Exhibit 5

Does High Frequency Market Manipulation Harm Market Quality?*

Jonathan Brogaard

David Eccles School of Business,
University of Utah

Dan Li

School of Management and Economics,
Chinese University of Hong Kong, Shenzhen

Jeffrey Yang

David Eccles School of Business,
University of Utah

April 2024

Abstract: Manipulation of financial markets has long been a concern. With the automation of financial markets, the potential for high frequency market manipulation has arisen. Yet, such behavior is hidden within vast sums of order book data, making it difficult to define and to detect. We develop a tangible definition of one type of manipulation, spoofing. Using proprietary user-level identified order book data, we show the determinants of spoofing. Exploiting lagged spoofing profitability and SEC Litigation Releases as instruments, we show causal evidence that spoofing increases volatility and adverse selection, and decreases price efficiency. The findings indicate that spoofing harms market quality.

JEL classification: G10, G12, G14
Keywords: high-frequency trading, market quality, market manipulation

*We thank Larry Glosten (discussant), Pete Kyle (discussant), Albert Lee (discussant), Gideon Saar (discussant), Sheri Belcher, Mike Cooper, Davidson Heath, Sam Draddy, and participants at the 18th Central Bank Conference on the Microstructure of Financial Markets, 18th Annual Conference on Asia-Pacific Financial Markets, City University of Hong Kong, Chinese University of Hong Kong, Finance Down Under, the NBER SI Big Data and High-Performance Computing for Financial Economics Conference, University of Sydney, University of Utah, and University of Western Ontario for comments. Send correspondence to Jonathan Brogaard, brogaardj@eccles.utah.edu.

Modern financial markets are largely automated. With the increased automation, market participants can potentially distort markets to profitably induce short term price movements. One such high-frequency manipulation method is spoofing, which is defined as "bidding or offering with the intent to cancel the bid or offer before execution."[1] In September 2020, JPMorgan was fined $920 million for spoofing metals and U.S. Treasury futures, where it was suggested that spoofing is a common practice.[2,3] The frequency of spoofing activity in financial markets is an empirical question. In addition, the fact that spoofing should be unrelated to real information and therefore does not contribute to price discovery raises the question of how spoofing affects market quality. This paper quantifies the frequency of spoofing and tests whether it harms market quality.

Theory on the impact market manipulation should have on market quality is mixed. Skrzypacz and Williams (2021) address the determinants and market quality impacts of spoofing. They theoretically show that increased spoofing activity leads to slower price discovery, higher return volatility, and wider posted and executed bid-ask spreads. A spoofing strategy impedes price discovery by driving prices away from fundamental values. Because deviations from fundamentals can be corrected, spoofing price movements induce reversals which then increase return volatility. At the same time, if spoofing drives prices away from fundamentals, adverse selection increases and market-makers are forced to raise spreads to remain profitable.

Some theoretical work argues against manipulation being feasible or that it can even improve market quality. Jarrow (1992) shows that when prices do not exhibit momentum, manipulation is not possible. Cherian and Jarrow (1995) show that a symmetric price response to manipulation renders it unprofitable. Other studies show that manipulation may be associated with improved market quality. Hanson and Oprea (2009) model a manipulator as a noise trader and show that the manipulation strategy encourages information acquisition as the profits to informed traders increase, thereby improving price accuracy. We empirically test these conflicting theories on the existence and effect of market manipulation.

We study Canadian equity markets using a proprietary IIROC dataset, which has trade and quote data with trader identification. We conservatively identify potential spoofing orders by

[1] 2010 Dodd-Frank Act
[2] https://www.reuters.com/article/jp-morgan-spoofing-penalty-idINKBN26K325
[3] https://fortune.com/2022/07/20/former-jpmorgan-trader-reveals-how-his-mentor-taught-him-to-place-and-cancel-bogus-spoof-trades-manipulate-markets/

applying six tractable filters to the data. We then examine the prevalence and determinants of spoofing in Canadian equity markets. We find that the median stock-day observation has 25 attempted spoofing orders. We exploit variation in spoofing from lagged spoofing profitability and SEC Litigation Releases to estimate the causal effect of spoofing on market quality. Our results are generally consistent with the theoretical predictions of Skrzypacz and Williams (2021). Spoofing leads to higher return volatility, higher adverse selection, and slower price discovery. However, we do not find strong evidence that spoofing increases transaction costs.

To discourage spoofing activity regulators strategically make the definition ambiguous. While is it not possible to perfectly identify spoofing orders, we draw from recent spoofing court cases[4] to develop a conservative six-step filtering approach that identifies trade and order behavior consistent with spoofing. First, all spoofing orders are eventually deleted. Second, spoofing buy (sell) order prices must be greater (less) than or equal to one tick below (above) the prevailing NBB (NBO). We match potential spoofing orders to genuine orders, which are orders in the opposite direction from the same trader. Third, spoofing orders must be placed within one second of the genuine order. Fourth, the sum of spoofing order volume must be ten times larger than the genuine order volume (as one genuine order can have multiple corresponding spoofing orders). Fifth, the spoofing orders must be cancelled within one second after genuine orders are executed or cancelled. Lastly, we require that during the second a spoofing order is placed, the trader does not actually trade in the same direction as the spoofing order. As it is very challenging to empirically distinguish market making from spoofing manipulation, we purposely use strict criteria that can distinguish between the two. A limitation of such a strict definition is that we likely undercount the true spoofing activity.

We begin the empirical analysis by documenting the prevalence and determinants of spoofing activity. Plotting spoofing against lagged market quality characteristics, we find that spoofing is most prevalent in stock-days with intermediate levels of liquidity when measured with the effective spread, price impact, realized spread. However, spoofing is more common in periods with lower quoted spreads. Spoofing is also most prevalent given intermediate levels of return volatility and high levels of price efficiency when measured with the Hasbrouck (1993) pricing error σ. We also show that spoofing is highly profitable. We compare genuine order execution

[4] For example, *United States v. Coscia* and *United States v. Bases et al.*

prices with the NBBO and show that the average genuine order makes a 10 to 21 basis point profit relative to the NBBO.

Motivated by the theoretical predictions from Skrzypacz and Williams (2021), we next focus on the relation between spoofing and market quality. We estimate 30-minute OLS panel regressions of market quality measures on the attempted spoofing order volume scaled by trading volume, while controlling for lagged dollar spread, lagged price, lagged inverse price, lagged absolute return, lagged log of dollar volume, lagged Amihud (2002) illiquidity, and stock and 30-minute-date fixed effects. Spoofing is positively associated with 15-second, 1-minute, and 5-minute return volatility, variance ratios, and the Hasbrouck (1993) pricing error. We find that spoofing is negatively associated with the quoted spread.

There is a strong endogeneity problem. Spoofing traders likely endogenously select certain stocks and dates to spoof. For instance, Skrzypacz and Williams (2021) predict that spoofers endogenously choose to spoof when markets are not so illiquid that their spoofing orders can be identified by market makers but not so liquid that their spoofing orders are unable to move markets. We document a similar pattern. If spoofing activity is correlated with a stock's ex-ante liquidity, then our OLS estimates suffer from omitted variable bias, as ex-ante liquidity likely predicts market quality.

We propose two instruments to overcome the endogeneity concern. First, at the intraday level, we use the average spoofing profitability from the prior 30-minute interval to instrument for spoofing. Spoofing traders should tend to trade the stocks that they deem the most profitable to spoof, therefore previously profitable periods should be more likely to be spoofed in the future. However, we argue that past profitability does not plausibly affect current market quality through a channel other than by affecting the current spoofing level, after controls. The instrumental variables estimation shows that spoofing increases return volatility, adverse selection, variance ratios, and the Hasbrouck (1993) pricing error volatility. However, we do not find that the quoted and effective spread increase, as predicted by Skrzypacz and Williams (2021). The results provide evidence that spoofing harms market quality.

Second, at the stock-day level, we exploit SEC Litigation Releases as shocks to spoofing activity. We interpret market manipulation-related SEC Litigation Releases as positive shocks to the ex-ante legal risk of spoofing for stocks subject to SEC jurisdiction. In the three days after a

release, spoofing in US cross-listed stocks decreases relative to stocks that are only listed on Canadian exchanges. Because SEC Litigation Releases predict spoofing activity but likely do not affect market quality directly, we instrument for spoofing by comparing the effect of SEC Litigation Releases on US cross-listed and Canada only stocks. The results are similar to those in the intraday IV. At the lower frequency, spoofing increases volatility, effective spreads, and the Hasbrouck pricing error σ but has statistically weaker effects on the other liquidity measures and the variance ratio.

Finally, we conduct a variety of robustness tests. We re-estimate our intraday IV results using alternative definitions of spoofing, such as the successful and unsuccessful spoofing order volume. We also compare the effect of spoofing by HFTs and non-HFTs, and apply strict trader-level filters to the data. Across the varying robustness checks the results remain economically consistent.

This paper contributes to the extant literature on market manipulation (See Putnins, 2012 for a survey) and more specifically to the newer literature on high frequency market manipulation. There is a nascent theoretical literature on spoofing. In general, it is challenging to model limit order book dynamics (Parlour, 1998; Rosu, 2009). Theory has incorporated spoofing behavior into the equilibrium order book behavior. Skrzypacz and Williams (2021) provide an equilibrium model showing that spoofing behavior can harm liquidity, slow price discovery, and elevate volatility. Wang, Hoang, Vorobeychik, and Wellman (2021) also show that the presence of spoofers in an order book that is otherwise informative results in a decrease in investor welfare. Cartea, Jaimungal, and Wang (2020) model how spoofing can be used to increase an investor's revenue, and how potential legal fines can deter spoofing behavior. Using simulated limit order books, Withanawasam, Whigham, and Crack (2018) examine where manipulators may be more prevalent. Our study provides empirical tests of the theoretical implications of spoofing on market quality and confirms that spoofing harms market quality.

Legal scholars have argued more generally about the impact of spoofing. Fischel and Ross (1991) provide a framework for how the legal community analyzes manipulation in markets. They argue that it is difficult to identify manipulation without knowing trader intent. They propose that no trades should be considered manipulative, while behavior that gives a false sense of trading activity (i.e. wash trading or matched orders) is manipulative. McNamara (2016) tackles the ethical

and legal implications of high frequency trading, which covers spoofing and other limit order-based manipulation strategies. Miller and Shorter (2016) survey the literature on high frequency trading and market manipulation and discuss the regulatory and legislative reaction to crack down on behaviors such as spoofing. Canellos et al. (2016) provide an overview of spoofing cases that have occurred before and after Dodd-Frank. Fox, Glosten, and Guan (2021) provide a framework to consolidate the varying interpretations of what is and is not considered spoofing. Montgomery (2016) argues that spoofing may in fact improve the liquidity of financial markets. Dalko, Michael, and Wang (2020) argue that spoofing as a manipulative practice only arises because of behavioral biases of investors and microstructural systems.

The empirical work on spoofing is limited. The reason for the paucity of work on the topic is that it typically requires order book data with trader identifying information. That said, Tao, Day, Ling, and Drapeau (2022) have crafted a strategy to detect spoofing from public order books. Two other papers have identifying account information and study spoofing. Lee, Eom, and Park (2013) use data from Korea and show a positive correlation among spoofing and volatility and a negative correlation with market capitalization. Wang (2019) uses data from Taiwan futures and shows that spoofing is profitable and is correlated with higher volume, bid-ask spreads, and volatility. This paper makes two contributions to the empirical literature. First, we provide another tractable spoofing detection method that aims to be orthogonal to genuine market-making activity. Second, we are the first to provide causal evidence that spoofing negatively impacts market quality.

1. Data and Variable Construction

Our primary data source is the proprietary Investment Industry Regulatory Organization of Canada (IIROC) dataset. The data consists of trade and order data for 137 Canadian stocks from May 3, 2010 to July 19, 2011. The sample is a volume stratified sample of Toronto Stock Exchange (TSX) stocks plus the TSX60 index constituents. Penny stocks and stocks with less than 20 active days are excluded. 46% of the firms in the sample are cross-listed in the US. We observe trades and quotes on the Toronto Stock Exchange. We also observe Alternative Trading System (ATS) activity through the Alpha (ALF), Chi-X (CHX), Omega (OMG), Pure (PTX), and MATCH Now (TCM) platforms.

The trade and order data are timestamped at the 10-millisecond level and contain order submissions, amendments, cancellations, and executions. Importantly, trades and orders in the data have unmasked trader IDs that allow us to track individual trader positions and strategies across time. For each event, we observe trader ID, order ID, price, volume, NBB, NBO, exchange, and other information. Each order is assigned an order ID that can be used to track the status of an order over time. This is crucial for spoofing identification, as it allows us to track an individual trader's cancellations and amendments with precision. We require that each stock-day has at least \$1 million in trading volume to remove very illiquid stocks. For intraday analysis, we require that each stock-day-30-minute interval has nonzero trading volume. We drop observations with quoted spreads above 5% to remove potential data errors. We also ignore the top 95th percentile of the variance ratio and Hasbrouck (1993) pricing error σ, as the right tails have extreme outliers.[5]

1.1 Market Quality Measures

We construct liquidity and market quality measures from the IIROC data. We measure liquidity with time-weighted quoted spreads, volume-weighted effective spreads, volume-weighed realized spreads, volume-weighed price impact, and Amihud (2002) illiquidity. We measure volatility with 15-second, 1-minute, and 5-minute return volatility, and market quality with variance ratios and Hasbrouck (1993) pricing error σ.

We compute time-weighted quoted spreads by weighting $\frac{NBO-NBB}{NBBO\ midpoint}$ by the time each spread prevails for a given stock-day. We compute volume-weighted effective spreads by weighing $2 \times \frac{D_k(Price_k - NBBO\ midpoint_k)}{NBBO\ midpoint_k}$ by the volume at each trade, k, where D_k is a trade sign indicator equal to 1 if the trade was buyer initiated and -1 if the trade was seller initiated. To approximate liquidity provision revenue, we compute volume-weighted realized spreads by weighing $2 \times \frac{D_k(Price_k - NBBO\ midpoint_{k,t+5})}{NBBO\ midpoint_k}$ by the volume at each at each trade, k, where $NBBO\ midpoint_{k,t+5}$ is the NBBO midpoint five minutes after trade k. Price impact is computed as the difference between the effective spread and realized spread. Amihud (2002) illiquidity is

[5] More details about the IIROC dataset can be found in the internet appendix for *The Competitive Landscape of High-Frequency Trading Firms* by Boehmer, Li, and Saar (2018).

computed as the absolute value of daily returns divided by dollar volume for each stock day, multiplied by 10^6.

Return volatilities are computed at the 15-second, 1-minute, and 5-minute levels and are the standard deviation of returns using midquote prices. We compute Lo and MacKinlay (1988) variance ratios with 15-second and 5-minute return variances with $\left|1 - 20 \times \frac{Var_{15\,second}(ret)}{Var_{5\,minute}(ret)}\right|$. Lastly, we compute the Hasbrouck (1993) pricing error σ. Similar to Boehmer and Kelley (2009), we estimate the VAR system with five lags and include four variables: log midquote returns, trade sign indicator equal to 1 if the trade was buyer initiated and -1 if the trade was seller initiated, signed volume computed as the trade sign times the number of shares traded, and root signed volume computed as the trade sign times the square root of the number of shares traded. We set lagged variables to zero at the beginning of each day. Table 1 Panel A reports liquidity and market quality summary statistics at the 30-minute level, while Panel B reports liquidity and market quality summary statistics at the stock-day level.[6]

INSERT TABLE 1 ABOUT HERE

1.2 Spoofing Measures

As the official definition of spoofing is likely strategically ambiguous, it is difficult to empirically measure the prevalence of spoofing activity. We draw our criteria from the following example of a trader who successfully executes a sell spoofing strategy: suppose a trader wants to buy shares of a stock. The NBB and NBO are currently \$99 and \$100, respectively. The trader wants to buy at a price less than or equal to \$99 and will manipulate prices down. First, the trader places a buy order for the shares he wants to buy at \$98.75, which is less than the prevailing NBO. He then rapidly places a high-volume limit sell order at a price lower than \$100 (but higher than \$99 to avoid immediate execution) to mimic selling pressure. The market responds to the false selling pressure by adjusting the NBB and NBO down. However, the trader immediately cancels the limit

[6] The quoted and effective spreads in the intraday panel are larger than those in the daily panel, as we only require nonzero trading volume rather than defining a threshold in the 30-minute panel. The negative stock-day average realized spread is consistent with that in Malinova, Park, and Riordan (2013), who study a similar sample period using Canadian stocks.

sell order before it can be executed. Because the market responds to the selling pressure, the NBB decreases and falls below $98.75, which results in the trader's buy order executing. Figure 1 describes this strategy graphically.

INSERT FIGURE 1 ABOUT HERE

Our example yields a more general definition. A trader who is spoofing the market will initially place a bona fide "genuine" buy limit order at a price at or lower than the current best bid price. After placing the genuine order, the trader will enter "spoofing" sell orders that will create the impression that the market is facing selling pressure. This will drive prices down and lead to the genuine order being executed. Finally, the spoofer will cancel the spoofing sell order. The same story holds with genuine sell orders and spoofing buy orders. We develop six filters to classify orders as potential spoofing orders.

We separately identify buy and sell spoofing orders. We also require that spoofing activity occurs during the trading hours of 9:30 AM to 4 PM. We describe the procedure for identifying spoofing buy orders in detail.[7] The spoofing identification procedure relies on visible trader IDs to track spoofing strategies.

We first search for spoofing orders without considering the other side's genuine orders. The first filter requires that spoofing orders are eventually deleted. As spoofing strategies consist of rapid entrance and cancellation of orders in the same direction, we expect that a spoofer will cancel a vast majority of their spoofing orders. Our main spoofing detection strategy implicitly assumes that spoofing orders are not executed. Although it is likely that some spoofing orders are unintentionally executed, it is difficult to disentangle an executed spoofing order from a non-spoofing order. Second, if a spoofing order is to induce a market response, it must be somewhat aggressive. We require that buy spoofing order prices are greater than or equal to one tick below the previous NBB.

[7] The procedure to identify spoofing sell orders is nearly identical to the procedure used to identify buy orders. Switching "buy" with "sell" and changing the second filter to require that the spoofing sell order must be less than the NBO yields the spoofing sell order identification procedure.

We match each potential buy spoofing order to potential sell (genuine) orders from the same trader ID.[8] Our third criteria requires that spoofing orders occur within one second before or after the genuine order is placed, and that the genuine order is placed before the spoofing order is cancelled. This is consistent with a spoofing trader entering a reasonable genuine order and also spoofing the market to induce a price response. For there to be a price effect, spoofing orders again must be sufficiently aggressive. Our fourth filter captures this by requiring that for each genuine order, the sum of spoofing order volume must be ten times larger than the genuine order volume. Spoofing occurs at high frequencies. Our fifth and most aggressive filter requires that spoofing orders are cancelled within one second after genuine orders are either cancelled or executed. Lastly, our sixth filter requires that for a given spoofing buy order, the trader ID must not have executed a buy order in the same second. This is consistent with the one-sided nature of spoofing. If a trader is trying to manipulate prices in one direction, it is unlikely that they will trade in the direction of their spoofing orders (and if they did, then the spoofing strategy would be much less profitable).

We define four types of spoofing: successful, unsuccessful, losing, and attempted. Successful spoofing orders are spoofing orders with executed genuine orders, while unsuccessful spoofing orders have cancelled genuine orders. Losing spoofing orders are spoofing orders that are eventually executed. We remove the spoofing order cancellation and time filters to measure losing spoofs. For a given stock-day, we only measure losing spoofs from traders who place at least one attempted spoofing order. This case is an undesirable outcome for spoofers, who try to avoid execution of spoofing orders. Finally, attempted spoofing orders are the sum of the successful, unsuccessful, and losing spoofing orders. Our main measure of spoofing is the attempted spoofing order volume scaled by trading volume. Figure 2 provides a graphic that describes the different outcomes associated with the spoofing and genuine order for each measure.

INSERT FIGURE 2 ABOUT HERE

[8] Note that our matching procedure can match multiple spoofing orders to a single genuine order. Our spoofing detection algorithm can therefore also capture layering activity, which regulators often use interchangeably with spoofing. Layering can be viewed as spoofing, but with multiple non-bona fide orders at different prices.

Table 1 Panel C presents the 30-minute level summary statistics for spoofing activity. In our sample, the median 30-minute interval has 1 attempted spoofing order and 0 successful spoofing orders. Table 1 Panel D presents stock-day level summary statistics for spoofing. The median number of attempted spoofs for a stock-day is 25, while the median number of successful spoofs is 0. At both the daily and intraday levels, the median count of losing spoofs is 0, indicating that spoofers are generally successful at avoiding execution on the spoofing side. High frequency traders place around 75% of the sample's attempted spoofing orders. The number of observations for the HFT spoofing percentage in Panels C and D is low because the measure conditions on nonzero spoofing.

We explore the characteristics of spoofing traders. We classify a trader as a spoofer if they make at least one attempted spoofing order in the sample. For a given stock-day, the median spoofer places seven attempted spoofing orders. For a given trader-day, the median spoofer targets two stocks. 6% of the traders in the sample place at least one attempted spoofing order, while 4% place at least one successful spoofing order. Spoofers tend to be large traders. The mean non-spoofer trades $294 million in the sample period, while the mean spoofer trades $4.8 billion (which is greater than the 95th percentile of trader-volume). Conditioning on large traders who trade over $500 million in the sample, 29% of traders place at least one attempted spoofing order and 24% place at least one successful spoofing order.

Spoofing activity is right skewed, which suggests that spoofing may be heavily concentrated within certain time periods or stocks. We disaggregate successful and attempted spoofs into the buy and sell types and find that on average, buying spoofing activity is slightly more common than selling spoofing activity. This suggests that traders who wish to manipulate the market by spoofing tend to do so with upward price pressure.

1.3 Microstructure Controls

We compute average dollar spread, average price, and inverse price as microstructure controls in regression tests. Average price is computed as the dollar trading volume divided by share trading volume, and inverse price is equal to 1 divided by the average price. Average dollar spread is computed by multiplying the quoted spread by the average price.

2. Spoofing Activity

In this section, we first examine the conditions in which spoofing tends to be most prevalent using lagged market quality measures. We then turn to measuring the profitability of spoofing.

2.1 Determinants of Spoofing

We begin by examining the determinants of spoofing activity graphically. We compute the average number of attempted spoofing orders for 15 lagged market quality quantiles. Skrzypacz and Williams (2021) predict that spoofing activity should be most active in markets with moderate liquidity. We measure liquidity with quoted spread, effective spread, realized spread, and price impact. We also show the relation between spoofing and lagged volatility and price efficiency measures. The results are shown in Figure 3.

INSERT FIGURE 3 ABOUT HERE

Panel A shows that spoofing tends to occur in stocks with lower ex-ante quoted spreads. However, spoofing is most prevalent in stock-days with intermediate ex-ante effective spreads, price impact, and realized spreads. This is consistent with the Skrzypacz and Williams (2021) prediction that spoofing should be the most prevalent in markets with intermediate levels of liquidity, as spoofers target sufficiently liquid markets to avoid being caught, while targeting sufficiently illiquid markets to be able to effectively influence prices.

Panel B presents results for ex-ante volatility. Spoofing occurs the most in stock-days with moderate levels of intraday return volatility. Spoofing in periods of low return volatility may lead to a higher change of being caught, while spoofing in periods of high return volatility is less likely to move prices in the desired direction. Panel C shows that spoofing occurs the most in stocks with lower inverse market quality when measured with the Hasbrouck (1993) pricing error σ. That is, spoofing is more prevalent when prices are more efficient. This is likely because spoofers target periods where their spoofing orders are more likely to be falsely impounded into prices as new information, such as when algorithmic trading is prevalent. However, spoofing has no clear relation with the variance ratio.

We validate the spoofing measure by examining spoofing activity around the passage of

the Dodd-Frank act. Namely, we observe a decrease in spoofing in US cross-listed stocks relative to stocks that are only listed on Canadian exchanges because of the more stringent anti-fraud provisions in Dodd-Frank that only apply to US cross-listed stocks. Because only US cross-listed stocks are subject to US regulations, Dodd-Frank should not affect spoofing in Canada-only stocks. In untabulated results, we use a difference-in-difference approach where the treatment group is the set of US cross-listed stocks, and the time-series shock is the passage of Dodd-Frank. After Dodd-Frank is passed, the treatment group experiences a decline in spoofing relative to the control group. The results suggest that increases in the ex-ante legal risk of spoofing can deter spoofing activity. Furthermore, the results validate the spoofing measure. If the true level of spoofing falls because of Dodd-Frank, then a valid proxy for the true level of spoofing should also fall.

2.2 **Spoofing Profitability**

We next estimate the profitability of spoofing. Because spoofing orders push prices in directions favorable to genuine orders, it is extremely profitable. For example, Michael Coscia was able to profit $1.4 million through spoofing in a three-month period.[9] We measure spoofing profitability in two ways by comparing genuine order prices with the NBBO at the time the genuine order was placed.

Figure 4 shows the spoofing profitability measures in the case of a genuine sell order. We only consider the profitability of successful spoofing sequences, where the spoofing orders are not executed and the genuine order is executed. For the genuine sell (buy) order, our baseline measure of profitability is the difference between the sell (buy) price and the NBB (NBO) when the genuine order was placed. In other words, it compares the execution price between the executed genuine order and a market sell order. Our second more conservative measure is the difference between the sell (buy) price and the NBO (NBB). This measure compares the execution price between the executed genuine order and a limit sell order, assuming that the trader would have been able to have executed the limit sell order immediately. The conservative measure is always smaller than the baseline measure because the NBB is lower than the NBO.

INSERT FIGURE 4 ABOUT HERE

[9] https://www.fbi.gov/news/stories/trader-sentenced-in-spoofing-case-involving-market-manipulation

Table 2 presents summary statistics on the profitability of spoofing. For each stock-day, we compute the average profitability of spoofing in terms of dollar value and basis points. Note that the dollar profit does not consider the number of shares traded and is therefore on a per-share basis. The stock-day summary statistics show the average profitability for a single genuine order. On average, a genuine order has a profit of $0.03 (20.79 bp). The conservative measure also has a positive profit of $0.01 (10.12 bp). Across both the baseline and conservative measures, the results indicate that spoofing can be highly profitable when successful.

INSERT TABLE 2 ABOUT HERE

3. Intraday Spoofing and Market Quality

In this section we study the relation between spoofing and market quality at the intraday level. We begin by estimating OLS associations, then turn to the causal relation between intraday spoofing and market quality by exploiting variation in spoofing driven by lagged spoofing profitability.

3.1 OLS Relation

Guided by the theoretical predictions in Skrzypacz and Williams (2021), we examine the relation between spoofing activity and market quality. Namely, increased spoofing activity should be associated with higher return volatility, higher bid-ask spreads, and slower price discovery. We measure return volatility with 15-second, 1-minute, and 5-minute return volatility. We measure spreads with time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, and volume-weighted price impact. We measure price discovery with the variance ratio and Hasbrouck (1993) pricing error σ. For ease of interpretation and to reduce the effect of outliers, we take the natural log of the volatility, variance ratio, and Hasbrouck (1993) pricing error σ meaures. We measure spoofing at 30-minute intervals as the sum of attempted spoofing order volume scaled by trading volume. We standardize the spoofing measure for ease of interpretation. To estimate the intraday relation between spoofing and market quality, we estimate the following regression equation for each market quality measure:

$$Market\ Quality_{i,t,j} = \beta_1 Attempted\ Spoofing_{i,t,j} + \boldsymbol{\beta}X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j}$$

Where $Attempted\ Spoofing_{i,t,j}$ is the standardized attempted spoofing order volume scaled by trading volume for stock i on day t during 30-minute interval j, and X is a vector of controls that includes the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock fixed effects with ζ_i and 30-minute interval fixed effects with $\phi_{t,j}$, respectively. Results are presented in Table 3.

INSERT TABLE 3 ABOUT HERE

We scale spoofing order volume by trading volume to more easily compare across stock-days with different levels of trading activity, as 500 shares of spoofing volume may have a different effect on market quality in a stock-day with 1 million vs 10 million shares of trading volume. Because the spoofing variable is standardized, the interpretation of β_1 is that a one standard-deviation increase in scaled attempted spoofing orders is associated with a β_1 unit change in the dependent variable.

We apply an aggressive set of controls to help mitigate several endogeneity concerns. However, we acknowledge that the results are not causal. We first include within-stock-day controls to help alleviate concerns that spoofers may target periods with different levels of trading volume, return magnitude, liquidity, or ex-ante microstructure characteristics. We also include stock and date-30-minute interval fixed effects.

We include several microstructure and liquidity controls because the decision to spoof likely depends on a stock's ex-ante level of market quality (as shown in Figure 3). We include the lag of average price, inverse price, and average dollar spread. This is because spoofing may be easier to implement in stocks that are less tick constrained. Our controls for log dollar volume and Amihud (2002) illiquidity help control for liquidity, while the absolute return control alleviates concerns that spoofing traders might tend to target stocks with high or low return magnitudes. For interpretability, we standardize all control variables in all specifications. We lag the intraday control variables to avoid controlling for a downstream affect, as spoofing may, for example, also directly affect the current price and dollar spread (Angrist and Pischke 2009).

The stock fixed effects sweep out any time-invariant and firm specific characteristics. The coefficient on spoofing is therefore a within-stock estimate of spoofing on market quality. The fixed effects can alleviate concerns that stock-specific confounders affect both spoofing and market quality. For example, the fixed effects sweep out potential endogeneity driven by average levels of firm liquidity, price efficiency, or spoofing profitability.

The date-30-minute interval fixed effects mitigate concerns that spoofing and market quality may be higher or lower during different times (Lee, Eom, and Park 2013). For example, a within stock analysis may be confounded by the fact that markets face more spoofing and higher volatility during the open and close, which would drive a positive relation between spoofing and volatility. The date-30-minute interval fixed effects alleviate this concern because they remove the time-series effect on both market quality and spoofing. It also removes potential intraday market wide shocks that affect both market quality and spoofing.

Thus, potential confounders would have to be time varying within stock, orthogonal to the intraday controls, orthogonal to time fixed effects, and affect both the spoofing and market quality measures. The aggressive fixed effects allow us to examine the effects of spoofing on market quality while mitigating many potential endogeneity concerns. However, because we do not exploit exogenous variation in intraday spoofing, we do not make causal claims in this section.

An increase in intraday spoofing is associated with significant increases in return volatility. Because we take the log of volatility, the coefficients on spoofing have a log-linear interpretation. A one standard-deviation increase in 30-minute spoofing is associated with a 4% increase in 15-second volatility, 3% increase in 1-minute volatility, and a 2% increase in 5-minute volatility. Spoofing has no clear association with spreads. Finally, intraday spoofing is associated with impeded price efficiency. A one standard-deviation increase in spoofing is associated with a 5% increase in the variance ratio and 11% increase in the Hasbrouck (1993) pricing error σ. The results suggest that spoofing is associated with worsened market quality at the intraday level.

A potential shortcoming in our spoofing identification approach is that we cannot determine a trader's true intent and thus may be instead measuring genuine market-making activity. It is unlikely that genuine market-making activity will manifest in our measures because of our sixth filter: a trader must not place a spoofing order in the same second that they trade in that direction. Our sixth filter likely removes much market-making activity as market-making

liquidity providers are more likely (or are required) to have balanced strategies. For example, the TSX appoints market makers who are required to maintain a two-sided market.

Although we control for likely confounders and include granular fixed effects, it is possible that omitted variables or reverse causality may bias our estimates. Thus, the results in this section can be viewed as associations between spoofing and market quality and are largely consistent with existing theoretical and empirical studies. Our finding that effective and realized spreads widen is consistent with Wang (2019), and the finding that return volatility is higher is consistent with Lee, Eom, and Park (2013). However, to our knowledge, we are the first to relate spoofing activity directly to price discovery measures such as variance ratios and Hasbrouck (1993) pricing errors.

3.2 Lagged Spoofing Profitability IV

The results in Table 3 may suffer from omitted variable bias or reverse causality, as it is likely that spoofing traders endogenously respond to current liquidity or market quality conditions that may make spoofing strategies more profitable or effective. We instrument for intraday spoofing using variation from the lagged profitability of spoofing. Because the choice to spoof likely depends on the ex-ante level of profitability, we argue that lagged profitability should have a positive relation with spoofing, as spoofers target periods that are more likely to be profitable.

We measure the profitability of spoofing using the measure developed in Section 2.2, which compares the execution price of genuine orders with the current NBBO. In particular, a genuine buy (sell) order is compared with the NBO (NBB) at the time of the buy (sell) order and indicates how much better off the spoofer was relative to placing a buy (sell) market order at the time the genuine order was placed.

From conversations with regulators, we learn that spoofers typically spoof securities when they receive a sufficient response from the market. In particular, spoofers will stop spoofing when a security becomes less profitable to spoof. This could occur for several reasons. For example, the algorithm that the spoofer profits off of may stop responding to the spoofing orders due to hitting loss limits, or a spoofing victim changes their algorithm in response to the losses from spoofing. In each case, the spoofer will spoof until spoofing becomes unprofitable. We measure the loss of profitability using the lagged gains from spoofing. Higher levels of profitability predict higher spoofing levels, while lower levels of profitability predict lower or zero spoofing levels.

We exploit the variation in spoofing using an instrumental variable (IV) approach, where the first stage regresses spoofing on lagged profitability in the following equation:

$$Attempted\ Spoofing_{i,t,j} = \beta_1 Spoofing\ Gain_{i,t,j-1} + \boldsymbol{\beta} X + \zeta_i + \ \phi_{t,j} + \epsilon_{i,t,j}$$

Where $Attempted\ Spoofing_{i,t,j}$ is the 30-minute standardized attempted spoofing order volume scaled by trading volume and $Spoofing\ Gain_{i,t,j-1}$ is the standardized average dollar gain to spoofing in previous 30-minute interval. X again represents lagged controls for the average dollar spread, average price, inverse price, log of dollar volume, absolute return, and Amihud (2002) illiquidity. We include stock and 30-minute fixed-effects with ζ_i and $\phi_{t,j}$, respectively. The first-stage results are presented in Table 4. The instrumental variable is only valid if it satisfied both the relevance and exclusion restrictions.

INSERT TABLE 4 ABOUT HERE

The first-stage result in Table 4 shows a strong and positive relation between contemporaneous spoofing and lagged spoofing gain, suggesting that the first stage is powerful. The positive relation is consistent with the hypothesis that higher profits to spoofing leads to more future spoofing, while lower profits to spoofing lead to less future spoofing. Because lagged spoofing gain is standardized, the coefficient represents that a one-standard deviation increase in the lagged spoofing gain leads to a 0.11 standard deviation increase in attempted spoofing. The T-statistic on $Spoofing\ Gain_{i,t,j-1}$ is 23.34 and the Kleibergen-Paap F-statistic (shown in Table 5) is greater than 438, which suggests that the first stage is sufficiently strong and therefore satisfies the relevance condition. We report the Kleibergen-Paap F-statistic in the second-stage results because the dependent variables have slightly different sets of non-missing observations.

The exclusion restriction requires that lagged spoofing profitability affects current market quality only though affecting current levels of spoofing. We argue that this assumption is satisfied. Potential violations of the exclusion restriction would have to be correlated with both the instrument and market quality and orthogonal to the second-stage controls, which we believe to be unlikely. For example, a potential concern is that a stock's average level of price efficiency makes it easier to spoof, and that the instrument may be correlated with both the stock's average price

efficiency and spoofing. However, as the average level of price efficiency is time-invariant, this variation would be swept out by the stock fixed effects. In untabulated results, we address two potential concerns. First, lagged profitability may be driven by lagged market quality, which may also affect current market quality through autocorrelation. For each dependent variable, we test whether including a one-day lagged control for the dependent variable affects the results. The results, available upon request, are consistent with the main specification. Second, lagged profitability may also drive future liquidity provision. For instance, at high frequencies, highly profitable spoofing leads voluntary liquidity providers to reduce their liquidity provision, thereby harming market quality in the near future. To address this concern, we develop a measure of market making by directly identifying stock-day market makers in the data using a list of market maker and stock-day pairs. We show that controlling for market making (market making order volume scaled by total order volume) does not change the results.

The second stage IV estimates are shown in Table 5. We estimate equations of the following form:

$$Market\ Quality_{i,t,j} = \beta_1 \widehat{Attempted\ Spoofing}_{i,t,j} + \boldsymbol{\beta} X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j},$$

where the dependent variables of interest and controls are as defined in the first stage specification. $\widehat{Attempted\ Spoofing}_{i,t,j}$ is the predicted standardized attempted spoofing volume scaled by trading volume for stock i on day t from the first-stage IV regression in Table 4.

INSERT TABLE 5 ABOUT HERE

The results show that spoofing increases return volatility. A one standard-deviation increase in spoofing causes a 28%, 24%, and 19% increase in 15-second, 1-minute, and 5-minute return volatilities, respectively. This is consistent with the idea that spoofing can move prices. If a spoofing trader can induce a temporary mispricing, then the process of inducing and correcting the manipulation will mechanically cause return volatility to increase.

Spoofing has a mixed effect on liquidity. A one standard-deviation increase in spoofing causes a 2.22 basis point decrease in the quoted spread, which suggests that spoofing strategies tend to decrease the quoted spread through aggressive orders with prices within the bid-ask spread.

Although the quoted spread decreases with spoofing, the orders that narrow the spread are likely non-bona fide orders that are never executed. However, the effect of spoofing on the effective spread is positive but statistically insignificant. A one standard-deviation increase in spoofing causes a 0.39 basis point increase in the effective spread. Spoofing also affects the realized spread and price impact. A one standard-deviation increase in spoofing leads to a 1.24 basis point decrease in the realized spread, consistent with spoofing strategies leading to a decline in market making profitability. Because spoofing also heightens the adverse selection problem, the price impact measure also increases. The results are consistent with the Skryzpacz and Williams (2021) argument that spoofing magnifies the adverse selection problem. However, spoofing compresses the quoted spread and has a statistically weak relation with the effective spread.

Spoofing causes higher variance ratios and Hasbrouck σ. A one standard-deviation increase in spoofing leads to a 17% increase in the variance ratio and a 5% increase in Hasbrouck σ, which is evidence that spoofing harms price discovery. As the variance ratio measure increases, the ratio of twenty 15-second return variances and 5-minute return variance deviates more from 1. This is evidence that increased spoofing activity drives price movements away from a random walk process, which indicates impeded price efficiency. The Hasbrouck (1993) procedure decomposes stock prices into random walk (efficient) and stationary (pricing error) components. Hasbrouck σ measures the variance of the pricing errors. Larger dispersion in pricing errors suggests a less efficient price process that tends to deviate more from true prices. Thus, the Hasbrouck σ result suggests that spoofing also causes lower price efficiency.

4. Relation between Daily Spoofing and Market Quality

We turn to examine the relation between spoofing and market quality at the stock-day level. We again begin with OLS correlations and then study the causal effect at the stock-day level using SEC Litigation releases in an IV framework.

4.1 OLS Relation

We begin by estimating stock-day OLS regressions of market quality in spoofing. For each market quality measure, we estimate regressions of the following form:

$$Market\ Quality_{i,t} = \beta_1 Attempted\ Spoofing_{i,t} + \boldsymbol{\beta} X + \gamma_t + \zeta_i + \epsilon_{i,t}$$

Where $Attempted\ Spoofing_{i,t}$ is the standardized attempted spoofing order volume scaled by trading volume, and X is a vector of controls that includes lagged average dollar spread, lag average price, lag inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We denote date and stock fixed effects with γ_t and ζ_i, respectively.

Similar to the intraday results, the stock-specific and time-varying controls help control for liquidity, microstructure characteristics, or market conditions. However, because spoofing happens at high frequencies, we are able to use contemporaneous controls for absolute return, log of trading volume, and Amihud (2002) illiquidity to better explain variation in the market quality variables. This is because the daily specifications likely do not suffer from the potential downstream control problem (Angrist and Pischke 2009). However, we lag the microstructure variables as spoofing may still directly affect contemporaneous microstructure controls. Stock fixed effects sweep out time-invariant stock-specific variation, such as industry or average levels of liquidity. Day fixed effects sweep out marketwide time variation, such as marketwide liquidity shocks.

INSERT TABLE 6 ABOUT HERE

The results in Table 6 show a positive relation between spoofing activity and most of the inverse market quality measures. Spoofing has no clear correlation with return volatility at the stock-day level. Spoofing is positively associated with effective spread but is associated with decreased quoted spreads. A one standard-deviation increase in successful spoofing orders is associated with a 0.14 basis point increase in the volume-weighted effective spread. However, the spoofing coefficients on the realized spread and price impact are statistically insignificant. On the other hand, a one standard-deviation increase in successful spoofing orders is associated with a 0.17 basis point decrease in the quoted spread. Lastly, spoofing is associated with slowed price discovery. A one standard-deviation increase in successful spoofing orders is associated with a 6% increase in the variance ratio and a 4% increase in the Hasbrouck (1993) pricing error σ.

4.2 SEC Litigation Releases IV

Because the intraday OLS results again suffer from a potential reverse causality problem, we turn to exploiting exogenous variation in spoofing to show a causal relation between spoofing and market quality. We exploit variation in spoofing induced by SEC litigation releases. The SEC issues litigation releases for its civil lawsuits in federal court. The press releases range from initial charges filed by the SEC to final judgement announcements. We focus specifically on market manipulation related press releases that occur in the sample as shocks to spoofing activity. SEC litigation releases have been studied in other papers such as Kacperczyk and Pagnotta (2023) and Aggarwal and Wu (2006).

SEC litigation releases likely affect the trading behavior of manipulative traders. We interpret litigation releases as positive shocks to the ex-ante legal risk of spoofing. Because regulators study limit order book data in market manipulation cases, a larger regulator presence increases the probability that manipulation is identified. If a spoofing trader observes that the SEC has begun or completed an investigation on market manipulation, the trader may infer heightened regulatory attention and thus a higher chance of being caught spoofing. The trader will thus reduce spoofing activity to reduce the chance of being caught.

We search the SEC Litigation Releases database for market manipulation releases.[10] A release is considered market manipulation if it contains the keyword "manipulation" and refers to stock price manipulation. For example, on September 24, 2010, the SEC charged four individuals with manipulating microcap stock prices. The traders allegedly engaged in a scheme to inflate two microcap stock prices and give a false sense of market liquidity in the stocks. Such events create a sense of heightened regulatory attention on market manipulation and should therefore discourage spoofing activity. We identify 21 SEC litigation releases on market manipulation in the sample period. To identify only the most severe shocks to the ex-ante legal risk of spoofing, we filter the list of releases to only include charges, allegations, sentences, and final judgements. The final list consists of 12 SEC releases. Figure 5 plots the litigation days in the sample.

INSERT FIGURE 5 ABOUT HERE

[10] https://www.sec.gov/litigation/litreleases.htm

Because we study the trading activity of cross-listed firms on Canadian exchanges, the analysis is only economically valid if SEC litigation releases can affect trading on Canadian markets. This is achieved through the Exchange Act of 1934's section on foreign securities exchanges.[11] Specifically, the provision on Foreign Securities Exchanges bans brokers and dealers from violating SEC regulations when trading on international exchanges if the stocks are "organized under the laws of" the United States. Because cross-listed stocks must comply with U.S. regulations, their stocks are likely protected from manipulation by U.S. and Canadian traders, even on Canadian exchanges. This is consistent with recent litigation. In *Harrington Global Opportunity Fund v CIBC World Markets Corporation,* U.S. and Canadian traders spoofed shares of Concordia International Corporation, a company cross listed in Canada (TSX) and the U.S. (NASDAQ), in 2016. The court acknowledged that a share of Concordia stock is the same whether it is traded on a U.S. or Canada exchange. Therefore, the court argued that it had jurisdiction over Canadian traders spoofing on Canadian exchanges because manipulating shares of Concordia would affect prices on NASDAQ.

We exploit the differential effect of SEC litigation releases on spoofing by comparing US cross-listed and Canada-only stocks. Because SEC litigation risk does not apply to Canada-only stocks, there should be a larger reduction in spoofing in US cross-listed stocks relative to Canada-only stocks. We use the differential effect of SEC litigation releases on spoofing in US cross-listed and Canada-only stocks to instrument for spoofing activity. This approach is similar to those of Hendershott, Jones, and Menkveld (2011) and Malceniece, Malcenieks, and Putniņš (2019), who use difference-in-differences regressions as the first stage of an IV.

Our first stage estimate is the difference-in-differences regression of the standardized attempted spoofing order volume scaled by trading volume on the interaction between $US\ Listed_i$, which is an indicator equal to 1 if the stock is cross-listed in the US, and $Litigation_t$, which is an indicator equal to 1 if day t is one to three days after a SEC litigation release on market manipulation. We choose a short period for $Litigation_t$ to avoid capturing slower moving reductions in manipulation which may plausibly improve market quality, such as insider trading or corporate misconduct. We include controls for lagged average dollar spread, lag average price, lag inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We also

[11] 15 U.S. Code § 78dd

include stock and date fixed effects and cluster standard errors by stock. The first-stage results are presented in Table 7. The instrument is valid if it satisfies both the relevance and exclusion restrictions.

INSERT TABLE 7 ABOUT HERE

The first-stage results in Table 7 show that the instrument is powerful. The coefficient on $US\ Listed_i \times Litigation_t$ shows that in the three days after a SEC litigation release, US cross-listed stocks experience a 0.11 standard deviation decline in spoofing relative to Canada-only stocks. This is consistent with the hypothesis that SEC litigation releases cause spoofing activity to decrease in US cross-listed stocks, as traders reduce their spoofing activity in response to heightened legal risk. The T-statistic on $SEC_t \times Treat_i$ is -3.23 and the Kleibergen-Paap rk Wald F statistic (shown in Table 8) is greater than 8.4. The highly significant coefficient on the instrument and reasonable Kleibergen-Paap rk Wald F statistic suggest that the relevance condition is satisfied. We again present the Kleibergen-Paap F-statistic in the second stage results because the value changes slightly when dependent variables have a different number of nonmissing observations. Figure 6 shows this relation graphically. For ease of comparison, we demean stock-day spoofing levels with stock fixed effects.

INSERT FIGURE 6 ABOUT HERE

The exclusion restriction requires that $US\ Listed_i \times Litigation_t$ only affects market quality through spoofing. Threats to exclusion would have to be correlated with both $US\ Listed_i \times Litigation_t$ and market quality and orthogonal to the second stage controls. While it cannot be empirically tested, it is challenging to think of alternative possible channels by which SEC litigation releases affect market quality other than through lowering market manipulation activity. One potential concern is that the IV affects high frequency market manipulation other than spoofing, such as short selling manipulation, settlement manipulation, and wash trading. To alleviate these concerns, we conduct untabulated robustness tests that suggest that the results are

not driven by short selling manipulation or settlement manipulation. In particular, we estimate the results across buy and sell spoofs only and also remove options settlement dates. The results remain robust. Furthermore, other types of manipulation such as wash trading create a false impression of liquidity. Therefore, if we observe that spoofing harms market quality, this would be despite any decreases in wash trading which may have otherwise improved short term market quality.

The second stage estimates are shown in Table 8. We regress the market quality measures from Table 6 on the predicted standardized spoofing values from the first stage estimate in Table 7. We again control for lagged average dollar spread, lag average price, lag inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We also include stock and date fixed effects and cluster standard errors by stock.

INSERT TABLE 8 ABOUT HERE

The results are consistent with the intraday IV results. Spoofing reliably increases intraday return volatility. A one standard-deviation increase in spoofing increases 15-second, 1-minute, and 5-minute return volatility. This is again evidence that spoofing can likely induce short term deviations from fundamentals, which then leads to a correction that subsequently increases return volatility. Spoofing has no statistically detectible relation with the quoted spread, realized spread, and price impact. However, the positive effect on the quoted spread is likely due to market makers quoting wider spreads in response to spoofing activity. The negative effect on the realized spread and positive effect on price impact suggest that at the stock-day level, market-makers lose profits from spoofing activity due to increases in adverse selection. Spoofing increases the effective spread, which is evidence that spoofing increases transaction costs in the market. Lastly, spoofing increases the variance ratio (although the coefficient is statistically insignificant) and Hasbrouck pricing error σ, which again shows that spoofing harms price efficiency.

5. Robustness

We apply a battery of robustness tests to ensure that our results are not driven by our choice of spoofing measure, different trader types, nor a lack of trader-level spoofing filters. We focus on the intraday analysis, as spoofing and its effects primarily occur at higher frequencies.

5.1 Alternative Spoofing Definitions

The main results measure spoofing as the standardized attempted spoofing volume scaled by trading volume. We show that the results are robust to alternate definitions of spoofing. Namely, we turn to successful and unsuccessful spoofs.

We define successful spoofing orders as attempted spoofing orders that also result in the genuine side being executed. We define unsuccessful spoofing orders as attempted spoofing orders that result in the genuine side being cancelled. We again scale each measure by trading volume and standardize the measure for interpretation. For each alternative spoofing measure, we re-estimate the intraday IV regressions.

The intraday IV results are robust across the two spoofing measures. The first stage is statistically strong. The coefficient on $t-1$ spoofing profitability is positive for both measures and yields T-statistics of 17.79 and 22.89 for successful and unsuccessful spoofs, respectively. Furthermore, the Kleibergen-Paap F-statistics are greater than 316 and 424 for successful and unsuccessful spoofs, respectively. Table 9 Panels A and B present the second-stage intraday results for successful and unsuccessful spoofing, respectively.

INSERT TABLE 9 ABOUT HERE

The results in Table 9 Panels A and B suggest that the spoofing measure is robust to alternative definitions of spoofing. Spoofing reliably increases return volatility, lowers the quoted spread, decreases the realized spread, increases the price impact, increases the variance ratio, and increases the Hasbrouck σ. However, we see that at the intraday level, the effect of successful spoofing on market quality is generally larger than that of unsuccessful spoofing. This suggests that successful spoofing strategies tend to have the largest adverse effect on market quality. For

example, a one standard-deviation increase in successful spoofing leads to a 33% increase in 15-second volatility, while a one standard-deviation increase in unsuccessful spoofing only leads to a 28% increase.

5.2 **HFT vs Non-HFT Spoofing**

We next examine whether spoofing from HFTs has a different effect than spoofing from non-HFTs. In practice, both types of spoofing have been known to occur (for example, the manual spoofing by clicking a computer mouse in the JPMorgan case[12] and the algorithm-based spoofing in *U.S. v. Coscia* (2017)). However, the academic literature primarily focuses on potential high-frequency manipulation from HFTs (e.g. Dalko and Wang 2020 and Gai, Yao, and Ye 2012), as HFTs are highly sophisticated and operate at high speeds. We provide another test on whether HFTs can act as manipulators and harm market quality. We also test whether non-HFT spoofing can affect market quality.

To distinguish between HFTs and non-HFTs, we identify HFTs using the approach outlined in Boehmer, Li, and Saar (2018). We estimate the intraday IV using spoofing measured from only HFTs and spoofing measured from non-HFTs. Table 10 presents the results.

INSERT TABLE 10 ABOUT HERE

The intraday IV results are presented in Panels A and B. The first stage results are strong. The coefficients on lagged profitability are positive, with T-statistics of 31.05 and 20.15 for HFT and non-HFT spoofing, respectively. The Kleibergen-Paap F-statistics are greater than 888 and 314 for HFT and non-HFT spoofing, respectively. The results suggest that the first stage is strong. The second-stage results show that both types of spoofing harm market quality. Spoofing reliably increases volatility, decreases the quoted spread, decreases the realized spread, increases the price impact, and increases the variance ratio and Hasbrouck σ. However, the coefficients on HFT

[12] https://www.bloomberg.com/news/articles/2022-07-19/jpmorgan-trader-spoofed-so-fast-colleagues-urged-ice-on-fingers

spoofing are generally similar in magnitude to non-HFT spoofing, suggesting that both types of spoofing have a similar effect on market quality.

5.3 Trader-Level Filters

The main measure of spoofing is based on matching potential spoofing orders to genuine orders. However, the algorithm does not consider trader-level characteristics. Based on recent court cases described in the Internet Appendix, we introduce two new filters and show that our results are robust to more stringent trader-level criteria.

In practice, spoofers tend to have very low fill rates relative to other traders, as the spoofing orders are intended to be cancelled. We therefore first require that the spoofing trader's trade-to-order ratio is sufficiently low. Empirically, we require that the trade-to-order ratio is less than the stock-day median value. For example, in *U.S. v. Coscia* (2017), the expert witness report stated that Michael Coscia had a much higher cancellation rate relative to other HFTs.

Spoofing orders tend to be large, while genuine orders tend to be small. This is consistent with the findings in *U.S. v. Coscia* (2017) and *SEC v. Lek Securities Corporation et al.* (2019) in that the spoofing side often uses large or numerous orders that are not intended to be executed, while the genuine side has a lower volume. The second filter requires that the trader's fraction of large orders is larger than the stock-day median value. We classify orders as large if they are greater than the 90th percentile of the stock-day's order size. This filter therefore reflects the size imbalance between spoofing and genuine orders.

We re-estimate the intraday IV approach to estimate the effect of spoofing on market quality at the intraday level. The results are shown in Table 11. We again see that spoofing, even with trader-level filters, harms market quality.

INSERT TABLE 11 ABOUT HERE

6. Conclusion

We document evidence of spoofing behavior in Canadian equity markets and provide causal evidence that spoofing harms market quality. Consistent with the theoretical predictions in Skrzypacz and Williams (2021), spoofing increases return volatility, increases adverse selection, and slows price discovery. However, we do not find strong evidence that spoofing leads to wider bid-ask spreads.

We develop a tractable six-step filtering process to identify spoofing orders and study the prevalence of spoofing. Consistent with Skrzypacz and Williams (2021), we show that spoofing activity is single-peaked in liquidity when measured with spreads and volatility.

OLS regressions show that on average, spoofing activity is associated with worse market quality. Using lagged spoofing profitability and SEC Litigation Releases as instruments, we exploit exogenous variation in spoofing with instrumental variables frameworks to provide causal evidence that spoofing harms market quality at both the intraday and daily level.

This paper makes two contributions to the literature. First, we provide a tractable spoofing detection method that aims to be orthogonal to genuine market-making activity. Second, motivated by the theoretical predictions in Skrzypacz and Williams (2021), we are the first to provide causal evidence that spoofing harms market quality.

References

Aggarwal, R. K., & Wu, G. (2006). Stock market manipulations. *Journal of Business*, 79(4), 1915-1953.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31-56.

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Boehmer, E., Li, D., & Saar, G. (2018). The competitive landscape of high-frequency trading firms. *Review of Financial Studies*, 31(6), 2227-2276.

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial Studies*, 27(8), 2267-2306.

Canellos, G. S., Rangwala, T. S., Walfish, D. R., Jou, J. K., & Palladino, S. L. (2016). The law surrounding spoofing in the derivatives and securities markets. In *FIA L&C Conference*.

Cartea, Á., Jaimungal, S., & Wang, Y. (2020). Spoofing and price manipulation in order-driven markets. *Applied Mathematical Finance*, 27(1-2), 67-98.

Cherian, J. A., & Jarrow, R. A. (1995). Market manipulation. *Handbooks in Operations Research and Management Science*, 9, 611-630.

Dalko, V., & Wang, M. H. (2020). High-frequency trading: Order-based innovation or manipulation? *Journal of Banking Regulation*, 21, 289-298.

Dalko, V., Michael, B., & Wang, M. (2020). Spoofing: effective market power building through perception alignment. *Studies in Economics and Finance*, 37(3), 497-511.

Fischel, D. R., & Ross, D. J. (1991). Should the law prohibit manipulation in financial markets. *Harvard Law Review*, 105, 503.

Fox, M. B., Glosten, L. R., & Guan, S. S. (2021). Spoofing and its Regulation. *Columbia Business Law Review*, Forthcoming.

Hanson, R., & Oprea, R. (2009). A manipulator can aid prediction market accuracy. *Economica*, 76(302), 304-314.

Hasbrouck, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *Review of Financial Studies*, 6(1), 191-212.

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity?. *Journal of Finance*, 66(1), 1-33.

Jarrow, R. A. (1992). Market manipulation, bubbles, corners, and short squeezes. *Journal of Financial and Quantitative Analysis*, 27(3), 311-336.

Kacperczyk, M., & Pagnotta, E. S. (2023). Legal risk and insider trading. *Journal of Finance*.

Lee, C. M., Mucklow, B., & Ready, M. J. (1993). Spreads, depths, and the impact of earnings information: An intraday analysis. *Review of Financial Studies*, 6(2), 345-374.

Lee, E. J., Eom, K. S., & Park, K. S. (2013). Microstructure-based manipulation: Strategic behavior and performance of spoofing traders. *Journal of Financial Markets*, 16(2), 227-252.

Malceniece, L., Malcenieks, K., & Putniņš, T. J. (2019). High frequency trading and comovement in financial markets. *Journal of Financial Economics*, 134(2), 381-399.

McNamara, S. (2016). The law and ethics of high-frequency trading. *Minnesota Journal of Law Science and Technology*, 17, 71.

Malinova, K., Park, A., & Riordan, R. (2013). Do retail traders suffer from high frequency traders? *Working Paper*.

Miller, R. S., & Shorter, G. (2016). High frequency trading: Overview of recent developments (Vol. 4). *Washington, DC: Congressional Research Service*.

Montgomery, J. D. (2016). Spoofing, market manipulation, and the limit-order book. *Working Paper*.

Lo, A. W., & MacKinlay, A. C. (1988). Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1), 41-66.

Parlour, C. A. (1998). Price dynamics in limit order markets. *Review of Financial Studies*, 11(4), 789-816.

Putniņš, T. J. (2012). Market manipulation: A survey. *Journal of Economic Surveys*, 26(5), 952-967.

Roşu, I. (2009). A dynamic model of the limit order book. *Review of Financial Studies*, 22(11), 4601-4641.

Tao, X., Day, A., Ling, L., & Drapeau, S. (2022). On detecting spoofing strategies in high-frequency trading. *Quantitative Finance*, 22(8), 1405-1425.

Wang, X., Hoang, C., Vorobeychik, Y., & Wellman, M. P. (2021). Spoofing the limit order book: A strategic agent-based analysis. *Games*, 12(2), 46.

Wang, Y, (2019). Strategic Spoofing Order Trading by Different Types of Investors in Taiwan Index Futures Market. *Journal of Financial Studies*, 27(1), 65- 104

Williams, B., & Skrzypacz, A. (2021). Spoofing in Equilibrium. *Working Paper.*

Withanawasam, R., Whigham, P., & Crack, T. F. (2018). Are Liquid or Illiquid Stocks More Easily Manipulated? The Impact of Manipulator Aggressiveness. *Working Paper*.

Ye, M., Yao, C., & Gai, J. (2019). The externalities of high frequency trading. *Working Paper*.

Appendix

Table A1. Variable definitions

Variable	Definition
Spoofing measures	
Attempted spoofs	Order volume from attempted spoofing orders as defined by the procedure in Section 1.2, scaled by trading volume. Includes both successful and unsuccessful spoofs, meaning that the associated genuine order does not have to be executed.
Successful spoofs	Order volume from successful spoofing orders as defined by the procedure in Section 1.2, scaled by trading volume. Includes only successful spoofs, meaning that the associated genuine order must be executed.
Unsuccessful spoofs	Order volume from unsuccessful spoofing orders as defined by the procedure in Section 1.2, scaled by trading volume. Includes only unsuccessful spoofs, meaning that the associated genuine order must be cancelled.
Losing spoofs	Order volume from losing spoofs as defined in Section 1.2. Includes only losing spoofs, which are spoofing orders that are executed instead of cancelled.
Percent HFT attempted spoofs	Percentage of attempted spoofing order volume (as defined in Section 1.2) that is placed by HFTs.
Market characteristics	
15-second return volatility	Standard deviation of 15-second midquote returns.
1-minute return volatility	Standard deviation of 1-minute midquote returns.
5-minute return volatility	Standard deviation of 5-minute midquote returns.
Quoted spread	Time-weighted quoted spread, where each quoted spread is $\frac{NBO-NBB}{NBBO\ midpoint}$.
Effective spread	Volume-weighted effective spread, where each effective spread is $2 \times \frac{D_k(Price_k - NBBO\ midpoint_k)}{NBBO\ midpoint}$. D_k is a trade sign indicator equal to 1 if the trade was buyer-initiated, and -1 if the trade was seller-initiated.
Realized spread	Volume-weighted realized spread, where each realized spread is $2 \times \frac{D_k(Price_k - NBBO\ midpoint_{k,t+5})}{NBBO\ midpoint_k}$. D_k is a trade sign indicator equal to 1 if the trade was buyer-initiated, and -1 if the trade was seller-initiated. $NBBO\ midpoint_{k,t+5}$ is the NBBO midpoint five minutes after trade k occurs.
Price impact	The difference between the effective spread and the realized spread.

Variance ratio	Lo and MacKinlay (1988) variance ratios using 15 second and 5-minute midquote return variances: $\left\|1 - 20 \times \frac{Var_{15-second}(ret)}{Var_{5\ minute}(ret)}\right\|$.
Hasbrouck σ	Standard deviation of pricing errors from VAR system with five lags and four variables: log midquote returns, trade sign indicator equal to 1 (-1) if the trading price is buyer (seller) initiated, signed volume computed as the trade sign times the number of shares traded, and root signed volume computed as the trade sign times the square root of the number of shares traded.
Dollar trading volume	Total trading volume.
Absolute return	Absolute value of return for the time interval.
Amihud (2002) illiquidity	Absolute value of return divided by dollar volume, multiplied by 10^6.
Microstructure Controls	
Average price	Dollar trading volume divided by share trading volume.
Inverse price	1 / Average price
Dollar spread	Average price $\times$ quoted spread

Figure 1: Spoofing Example

Figure 1 provides a graphical representation of the sell spoofing example described in Section 1.2.

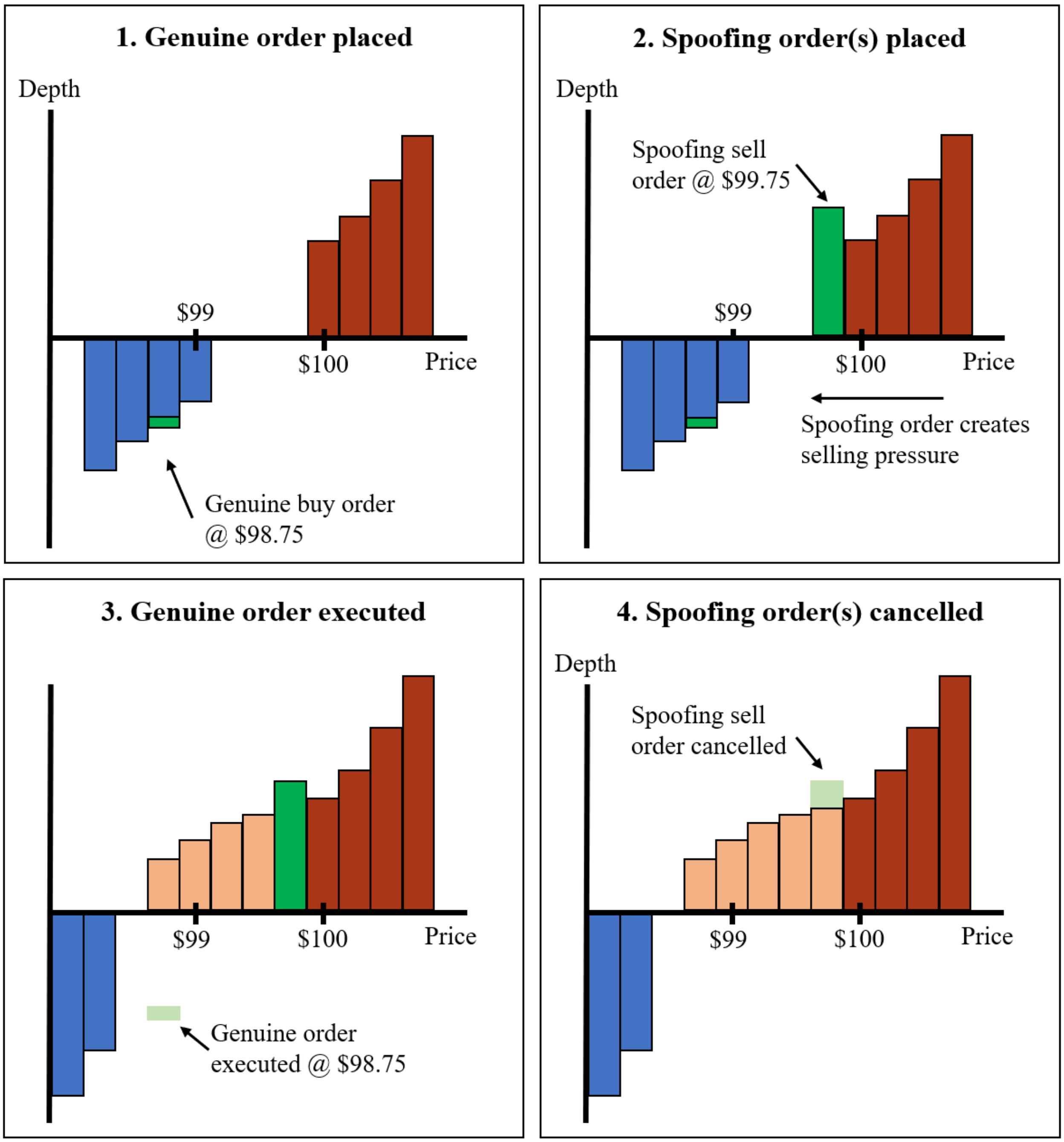


Figure 2: Spoofing Measures

Figure 2 describes the different types of spoofing measures developed in section 1.2. For each outcome of spoofing order and genuine order, the figure maps the pair to the corresponding spoofing definition.

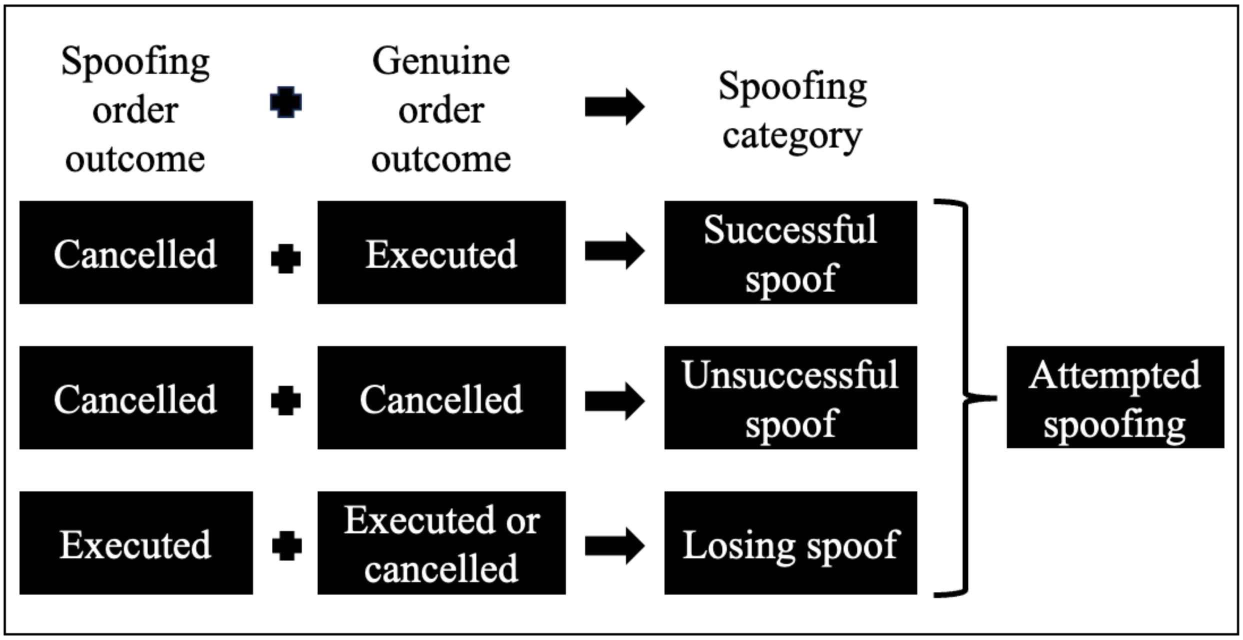
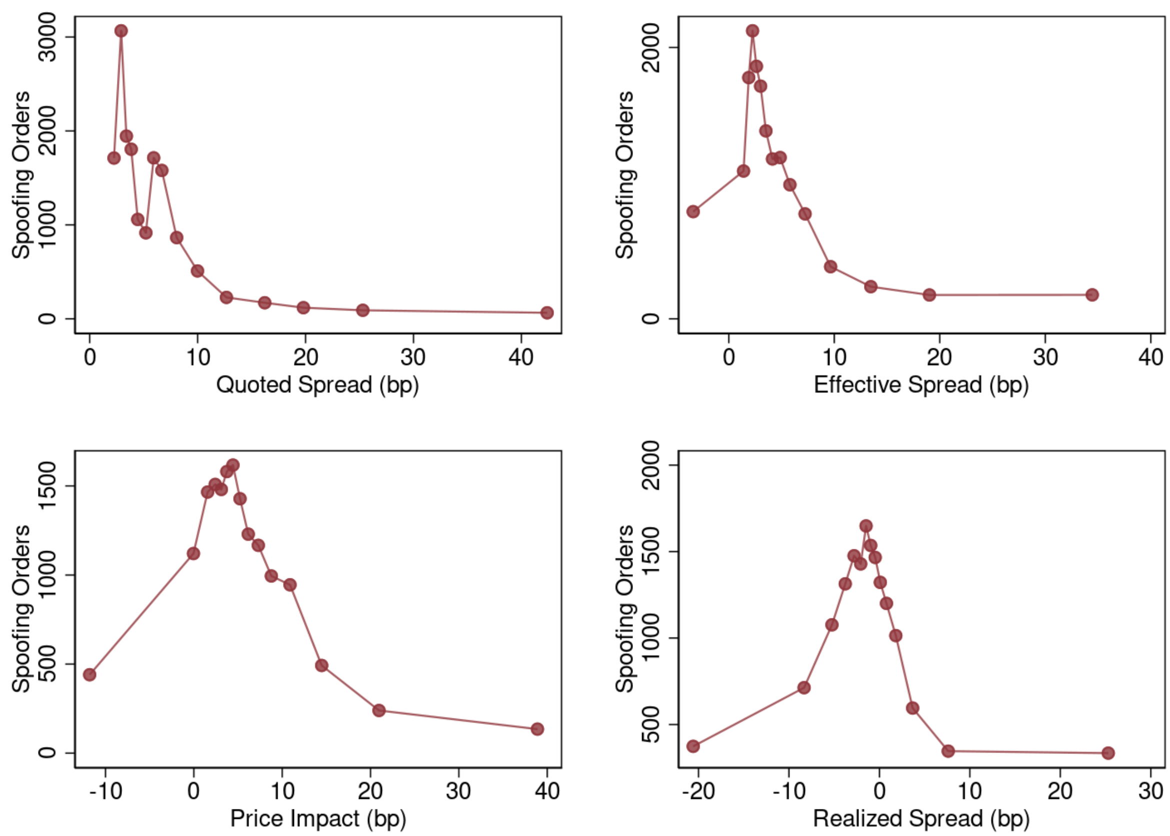


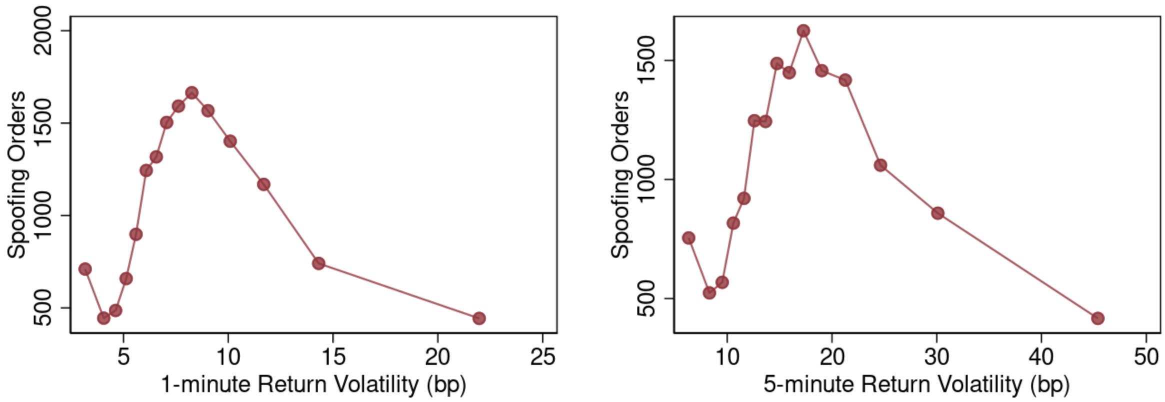
Figure 3: Spoofing and Market Quality

Figure 3 plots spoofing activity given lagged market quality quantiles. The vertical axis shows spoofing, measured as the number of attempted spoofing orders for a given stock-day. The horizontal axis represents different market quality measures. Panel A plots the average level of spoofing for each transaction cost quantile, where transaction costs are measured with time-weighted quoted spread, volume-weighted effective spread, volume-weighted price impact, and volume-weighted realized spread. Panel B plots the average level of spoofing given volatility quantiles, where volatility is measured with 1 and 5-minute return volatility. Panel C plots the average level of spoofing given price efficiency quantiles, where price efficiency is measured with the variance ratio and Hasbrouck (1993) pricing error σ.

Panel A: Spoofing and Lagged Transaction Costs



Panel B: Spoofing and Lagged Volatility



Panel C: Spoofing and Lagged Price Efficiency

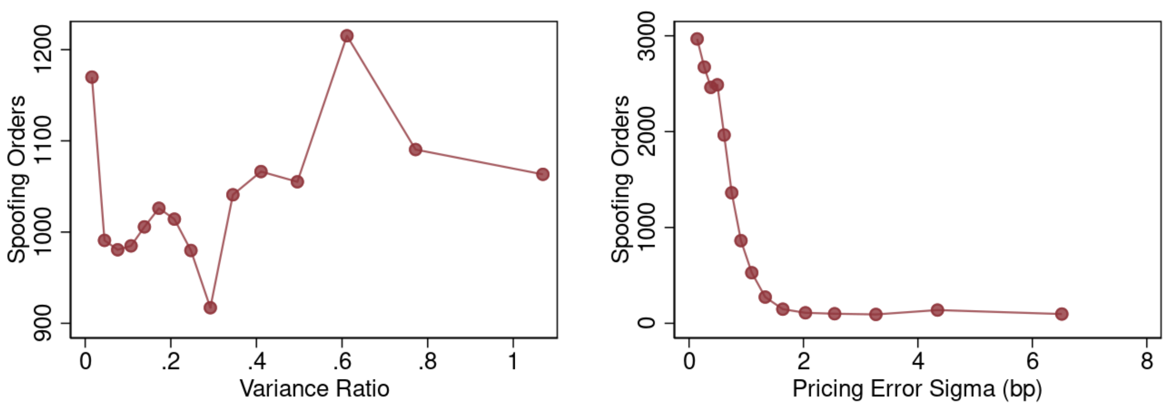


Figure 4: Spoofing Profitability

Figure 4 provides a graphical depiction of the spoofing profitability measures described in Section 2.2.

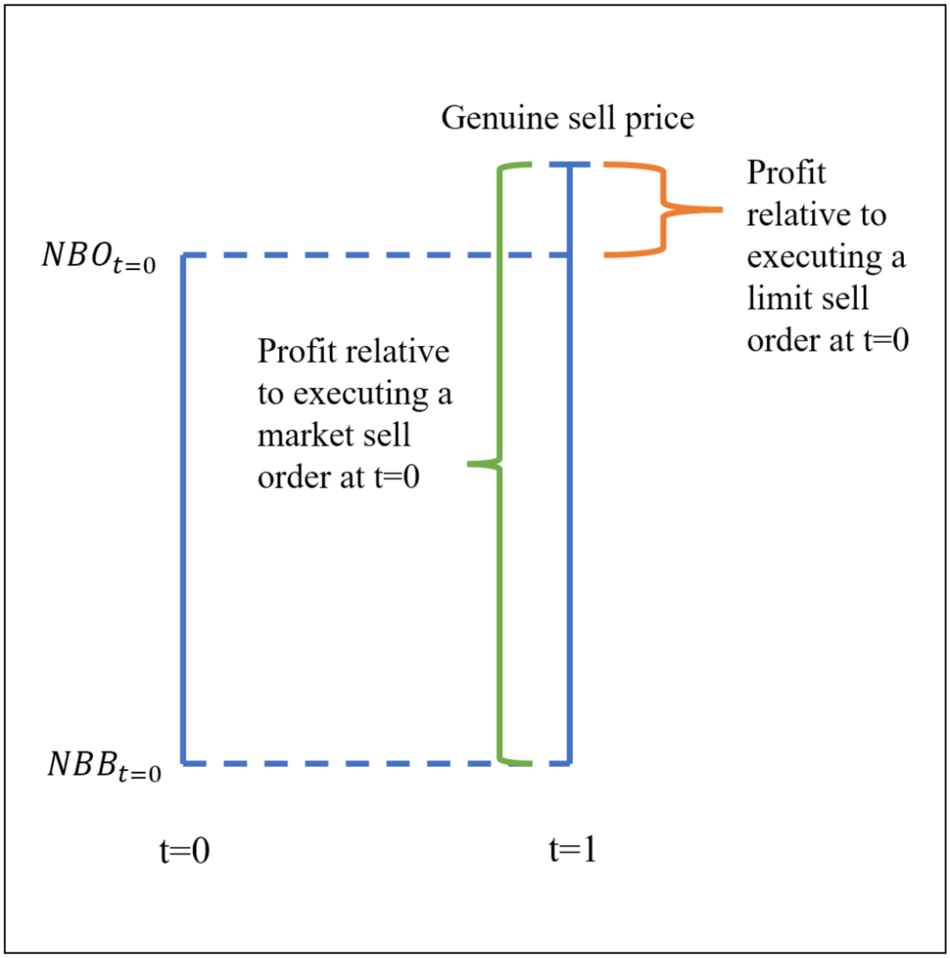


Figure 5: SEC Litigation Releases

Figure 5 plots the 12 SEC litigation releases in the sample period. The sample of SEC litigation releases consists of charges, allegations, sentences, and final judgements that are related to trade or order-based market manipulation. Litigation days are defined as those in the three days following an SEC litigation release.

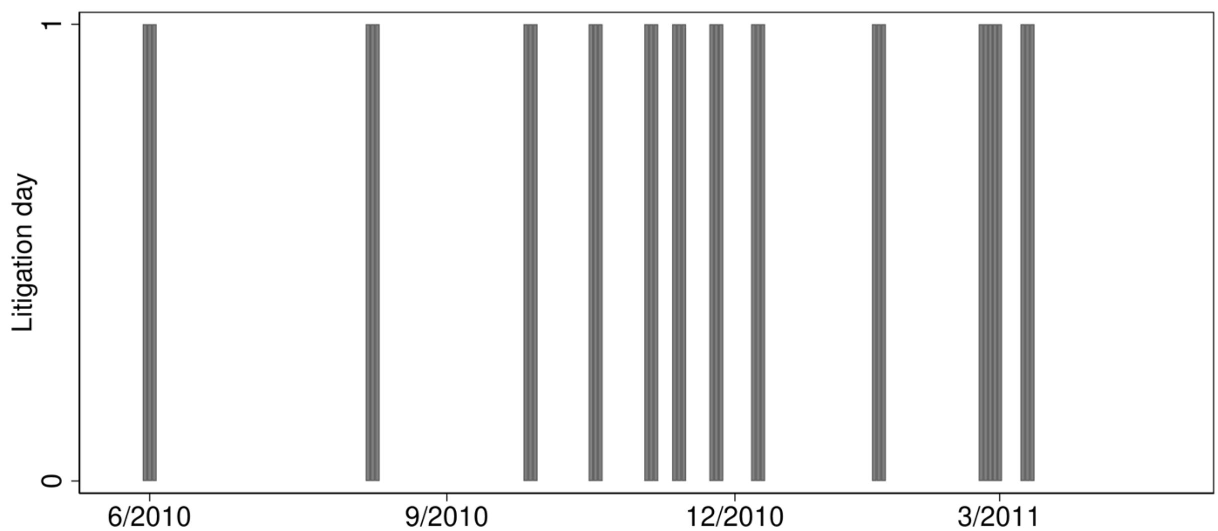


Figure 6: SEC Litigation Releases and Spoofing Activity

Figure 6 plots the average daily spoofing activity for US cross-listed and Canada-only stocks during litigation and non-litigation periods. Stock-day spoofing levels are demeaned with stock fixed-effects. Litigation periods are defined as the three days after a significant SEC litigation release on trade or order-based market manipulation. Spoofing is measured as the order volume from attempted spoofing scaled by trading volume.

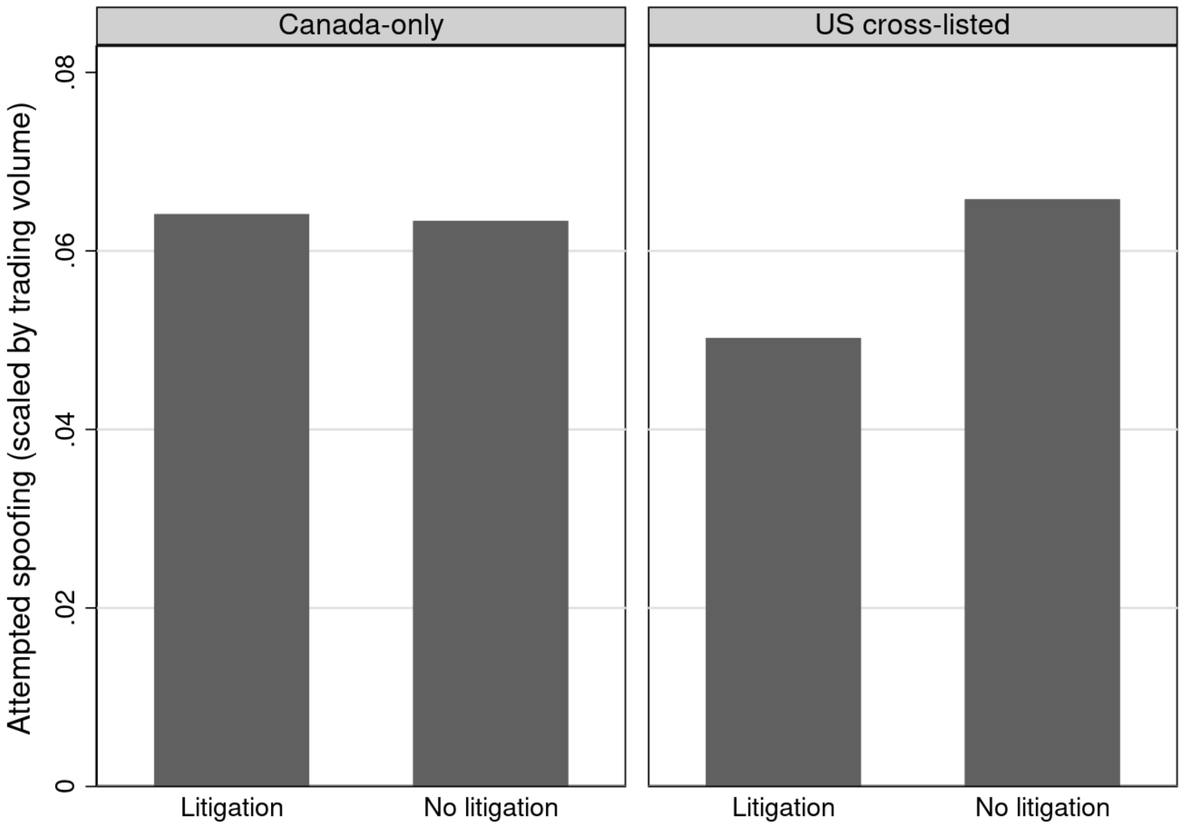


Table 1: Summary Statistics

Panel A presents 30-minute level summary statistics for market quality measures. Panel B presents stock-day level summary statistics for market quality measures. All Panel A and B variables except for the variance ratio are reported in basis points. Panel C presents 30-minute level summary statistics for spoofing activity. Panel D presents stock-day level summary statistics for spoofing activity. Spoofing variables as percentages of volume are defined as the spoofing order volume scaled by the level of total trading volume. Variables are winsorized at the 1% and 99% levels.

Panel A: 30-Minute Market Quality

	Mean	SD	p10	Median	p90	N
15-second volatility (bp)	4.55	4.15	1.29	3.3	9.09	346345
1-minute volatility (bp)	8.67	7.87	2.40	6.33	17.44	346075
5-minute volatility (bp)	17.51	16.58	4.11	12.52	36.3	344020
Quoted spread (bp)	32.23	56.17	3.04	11.8	80	369098
Effective spread (bp)	22.17	50.31	0.84	5.73	61.03	368980
Realized spread (bp)	10.42	48.75	-16.22	0.79	45.35	368980
Price impact (bp)	11.44	31.26	-4.12	4.1	36.23	368980
Variance ratio	0.84	1.08	0.06	0.42	2.31	322006
Hasbrouck σ (bp)	2.22	2.51	0.25	1.22	5.84	242254

Panel B: Stock-Day Market Quality

	Mean	SD	p10	Median	p90	N
15-second volatility (bp)	4.49	2.75	2.12	3.72	7.78	21413
1-minute volatility (bp)	8.59	5.04	4.10	7.18	14.84	21413
5-minute volatility (bp)	17.88	10.52	8.38	15	31.21	21413
Quoted spread (bp)	12.73	13.26	2.97	7.1	28.55	21413
Effective spread (bp)	8.12	10.85	1.43	4.35	20.96	21413
Realized spread (bp)	-0.15	11.54	-8.68	-0.93	8.52	21413
Price impact (bp)	8.21	12.73	-0.02	5.36	22.66	21413
Variance ratio	0.33	0.29	0.04	0.25	0.77	20326
Hasbrouck σ (bp)	1.86	1.84	0.28	1.15	4.64	20190

Panel C: 30-Minute Spoofing

	Mean	SD	p10	Median	p90	N
Attempted spoofs (% volume)	16.41	45.74	0.00	0.51	41.03	369098
Successful spoofs (% volume)	5.38	15.85	0.00	0.00	13.01	369098
Unsuccessful spoofs (% volume)	10.86	30.83	0.00	0.19	27.08	369098
Losing spoofs (% volume)	0.16	0.63	0.00	0.00	0.26	369098
Attempted spoofs (#)	66.39	263.2	0	1	144	369098
Attempted buy spoofs (#)	33.74	142.18	0	0	73	369098
Attempted sell spoofs (#)	32.65	148.94	0	0	65	369098
Successful spoofs (#)	1.51	5.85	0	0	4	369098
Unsuccessful spoofs (#)	62.61	257.04	0	1	131	369098
Losing spoofs (#)	2.26	10.67	0	0	3	369098
Percent HFT attempted spoofs (%)	75.55	35.51	0.00	97.07	100	115972

Panel D: Stock-Day Spoofing

	Mean	SD	p10	Median	p90	N
Attempted spoofs (% volume)	6.47	15	0.00	0.64	18.73	21413
Successful spoofs (% volume)	0.26	0.5	0.00	0.00	0.89	21413
Unsuccessful spoofs (% volume)	5.99	14.53	0.00	0.47	17.2	21413
Losing spoofs (% volume)	0.22	0.51	0.00	0.00	0.77	21413
Attempted spoofs (#)	1071.39	2786.41	0	25	3472	21413
Attempted buy spoofs (#)	545.69	1421.94	0	12	1759	21413
Attempted sell spoofs (#)	525.69	1408.45	0	9	1642	21413
Successful spoofs (#)	51.19	127.07	0	0	163	21413
Unsuccessful spoofs (#)	983.8	2641.78	0	19	3096	21413
Losing spoofs (#)	36.41	100.96	0	0	107	21413
Percent HFT attempted spoofs (%)	75.18	33.7	8.45	93.77	100	13888

Table 2: Profitability Per Share from Spoofing

Table 2 presents summary statistics for the profitability of spoofing strategies. We condition on spoofing sequences where the genuine order is executed and the spoofing order is not executed. For each stock-day observation, the spoofing gain is the average difference between the genuine order's execution price and the best bid or offer on the opposite side. The conservative spoofing gain is the average difference between the genuine order's execution price and the best bid or offer on the same side. The summary statistics are computed using stock-day averages of the profits to each genuine order.

	Mean	SD	p10	Median	p90	N
Spoofing gain ($)	0.03	0.22	0.01	0.02	0.03	10257
Spoofing gain (bp)	20.79	105.82	1.57	7.14	52.36	10257
Conservative spoofing gain ($)	0.01	0.28	-0.01	0.00	0.02	10257
Conservative spoofing gain (bp)	10.12	128.81	-3.03	1.42	25.91	10257

Table 3: Intraday Spoofing and Market Quality

Table 3 presents results of the following regression equation: $Market\ Quality_{i,t,j} = \beta_1 Attempted\ Spoofing_{i,t,j} + \boldsymbol{\beta}X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j}$, where $MarketQuality_{i,t,j}$ is log of 15-second return volatility, log of 1-minute return volatility, log of 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted price impact, log of variance ratio, or log of Hasbrouck (1993) pricing error σ for stock i on day t in the 30-minute interval j. $Attempted\ Spoofing_{i,t,j}$ is the standardized attempted spoofing order volume scaled by trading volume. X represents standardized $j-1$ controls for the average dollar spread, average price, inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock fixed effects with ζ_i and 30-minute time fixed effects with $\phi_{t,j}$. T-statistics are reported in parentheses and standard errors are clustered by stock-day.

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$Attempted\ Spoofing_{i,t,j}$	0.04***	0.03***	0.02***	-0.28***	-0.03	-0.05	0.02	0.05***	0.11***
	(22.39)	(15.05)	(8.20)	(-5.94)	(-0.44)	(-0.65)	(0.30)	(13.18)	(34.19)
$Average\ Dollar\ Spread_{i,t,j-1}$	0.13***	0.12***	0.12***	18.12***	10.29***	8.03***	1.89***	0.02***	0.22***
	(27.90)	(23.55)	(20.78)	(51.66)	(34.34)	(26.31)	(11.51)	(3.29)	(28.97)
$Average\ Price_{i,t,j-1}$	-0.09***	-0.09***	-0.10***	-7.57***	-4.89***	-3.38***	-1.29***	0.02	0.12***
	(-10.28)	(-10.93)	(-10.06)	(-24.45)	(-17.12)	(-11.27)	(-6.85)	(1.47)	(9.27)
$Inverse\ Price_{i,t,j-1}$	0.04***	0.05***	0.06***	31.33***	21.38***	18.03***	2.47***	-0.27***	0.34***
	(4.52)	(5.86)	(5.77)	(49.68)	(29.95)	(25.56)	(5.54)	(-17.02)	(21.41)
$\ln(Dollar\ Volume)_{i,t,j-1}$	0.05***	0.05***	0.05***	-0.16***	-0.13**	0.26***	-0.40***	-0.01***	-0.05***
	(16.59)	(16.06)	(15.69)	(-3.70)	(-2.34)	(4.43)	(-9.13)	(-2.92)	(-14.94)
$Absolute\ Return_{i,t,j-1}$	0.10***	0.10***	0.11***	2.99***	1.82***	-0.12	1.81***	0.02***	0.06***
	(42.40)	(42.18)	(40.40)	(23.53)	(10.65)	(-0.64)	(14.19)	(4.67)	(22.26)
$Amihud\ Illiquidity_{i,t,j-1}$	-0.02***	-0.01***	-0.01**	0.20	-0.05	0.13	-0.12	-0.01	0.07
	(-2.65)	(-2.74)	(-2.54)	(1.59)	(-0.28)	(0.79)	(-1.08)	(-0.85)	(1.10)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276,808	275,479	272,831	292,772	292,715	292,715	292,715	260,371	203,750
Adjusted R-squared	0.412	0.380	0.307	0.770	0.418	0.231	0.097	0.060	0.768

Table 4: First Stage Intraday IV Estimate

Table 4 presents results for the following regression equation: $Attempted\ Spoofing_{i,t,j} = \beta_1 Spoofing\ Gain_{i,t,j-1} + \boldsymbol{\beta}X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j}$, where $Attempted\ Spoofing_{i,t,j}$ is the standardized attempted spoofing order volume scaled by trading volume. $Spoofing\ Gain_{i,t,j-1}$ is the average dollar gain to spoofing, as defined in Section 2.2. X represents standardized $j-1$ controls for the average dollar spread, average price, inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock fixed effects with ζ_i and 30-minute interval fixed effects with $\phi_{t,j}$. T-statistics are reported in parentheses and standard errors are clustered by stock-day.

	(1) $Attempted\ Spoofing_{i,t,j}$
$Spoofing\ Gain_{i,t,j-1}$	0.11***
	(23.34)
$Average\ Dollar\ Spread_{i,t,j-1}$	0.02**
	(2.51)
$Average\ Price_{i,t,j-1}$	-0.23***
	(-8.17)
$Inverse\ Price_{i,t,j-1}$	-0.24***
	(-28.30)
$\ln(Dollar\ Volume)_{i,t,j-1}$	-0.05***
	(-8.35)
$Absolute\ Return_{i,t,j-1}$	-0.00
	(-0.94)
$Amihud\ Illiquidity_{i,t,j-1}$	0.00
	(0.30)
Stock FE	Yes
30-minute FE	Yes
Observations	292,772
Adjusted R-squared	0.447

Table 5: Second Stage Intraday IV Estimate

Table 5 presents results for the following regression equation $Market\ Quality_{i,t,j} = \beta_1 \widehat{Attempted\ Spoofing}_{i,t,j} + \boldsymbol{\beta}X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j}$, where $MarketQuality_{i,t,j}$ is the log of 15-second return volatility, log of 1-minute return volatility, log of 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted price impact, log of variance ratio, or log of Hasbrouck (1993) pricing error σ. $\widehat{Attempted\ Spoofing}_{i,t,j}$ is the predicted standardized attempted spoofing volume scaled by trading volume for stock i on day t and 30-minute interval j from the first-stage IV regression in Table 4. X represents standardized $j-1$ controls for the average dollar spread, average price, inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock fixed effects with ζ_i and 30-minute interval fixed effects with $\phi_{t,j}$. T-statistics are reported in parentheses and standard errors are clustered by stock-day.

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$\widehat{Attempted\ Spoofing}_{i,t,j}$	0.28***	0.24***	0.19***	-2.22***	0.39	-1.24**	1.50***	0.17***	0.05***
	(15.13)	(12.95)	(10.16)	(-9.27)	(0.75)	(-2.22)	(3.56)	(5.09)	(2.69)
$Average\ Dollar\ Spread_{i,t,j-1}$	0.13***	0.12***	0.11***	18.15***	10.29***	8.05***	1.87***	0.02***	0.22***
	(26.28)	(22.40)	(19.99)	(51.91)	(34.28)	(26.41)	(11.33)	(2.92)	(29.38)
$Average\ Price_{i,t,j-1}$	-0.03***	-0.05***	-0.06***	-8.03***	-4.79***	-3.67***	-0.94***	0.05***	0.12***
	(-2.77)	(-4.09)	(-4.96)	(-25.22)	(-15.18)	(-10.98)	(-4.33)	(2.91)	(8.94)
$Inverse\ Price_{i,t,j-1}$	0.11***	0.11***	0.11***	30.82***	21.49***	17.72***	2.86***	-0.23***	0.32***
	(10.23)	(10.73)	(9.59)	(47.75)	(28.92)	(24.01)	(6.14)	(-12.66)	(17.76)
$\ln(Dollar\ Volume)_{i,t,j-1}$	0.06***	0.06***	0.06***	-0.22***	-0.11*	0.22***	-0.35***	-0.01*	-0.06***
	(15.85)	(15.49)	(15.38)	(-4.67)	(-1.92)	(3.46)	(-7.56)	(-1.78)	(-15.18)
$Absolute\ Return_{i,t,j-1}$	0.10***	0.10***	0.11***	3.00***	1.82***	-0.11	1.80***	0.02***	0.07***
	(41.33)	(41.35)	(39.88)	(23.53)	(10.62)	(-0.61)	(14.11)	(4.50)	(22.28)
$Amihud\ Illiquidity_{i,t,j-1}$	-0.02***	-0.01***	-0.01**	0.20	-0.05	0.13	-0.12	-0.01	0.07
	(-2.64)	(-2.74)	(-2.54)	(1.59)	(-0.28)	(0.79)	(-1.08)	(-0.85)	(1.11)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276,808	275,479	272,831	292,772	292,715	292,715	292,715	260,371	203,750
Kleibergen-Paap F-statistic	455.4	454.7	453.4	544.6	544.6	544.6	544.6	456.7	438.6

Table 6: Stock-Day Spoofing and Market Quality

Table 6 presents results of the following regression equation: $Market\ Quality_{i,t} = \beta_1 Attempted\ Spoofing_{i,t} + \boldsymbol{\beta} X + \zeta_i + \gamma_t + \epsilon_{i,t}$, where $MarketQuality_{i,t}$ is the log of 15-second return volatility, log of 1-minute return volatility, log of 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted price impact, log of variance ratio, or log of Hasbrouck (1993) pricing error σ. $Attempted\ Spoofing_{i,t}$ is the standardized attempted spoofing order volume scaled by trading volume. X represents standardized controls for the lag of average dollar spread, lag of average price, lag of inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock and date fixed effects with ζ_i and γ_t , respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$Attempted\ Spoofing_{i,t}$	0.00	-0.00	-0.01	-0.17***	0.14**	0.10	0.07	0.06***	0.04***
	(0.30)	(-0.65)	(-1.44)	(-3.18)	(2.51)	(1.06)	(0.78)	(3.77)	(2.84)
$Average\ Dollar\ Spread_{i,t-1}$	0.09***	0.09***	0.09***	2.32***	1.22***	0.60***	0.72**	0.03	0.09***
	(8.18)	(8.82)	(9.01)	(6.42)	(4.92)	(2.83)	(2.44)	(0.84)	(3.01)
$Average\ Price_{i,t-1}$	-0.10***	-0.10***	-0.10***	-0.89*	-0.40	-0.30	0.04	-0.01	0.15
	(-3.10)	(-3.27)	(-3.34)	(-1.87)	(-1.16)	(-0.79)	(0.10)	(-0.21)	(1.49)
$Inverse\ Price_{i,t-1}$	0.06*	0.06**	0.07***	10.41***	6.27***	5.62***	1.12	-0.02	0.01
	(1.94)	(2.16)	(2.64)	(12.38)	(8.95)	(7.70)	(1.57)	(-0.42)	(0.11)
$\ln(Dollar\ Volume)_{i,t}$	0.23***	0.23***	0.22***	-0.43**	-2.68***	2.03***	-4.33***	0.01	-0.34***
	(14.96)	(14.71)	(13.37)	(-2.61)	(-9.24)	(3.35)	(-7.86)	(0.34)	(-11.88)
$Absolute\ Return_{i,t,j}$	0.06***	0.07***	0.08***	-0.09	0.41***	-0.95***	1.31***	-0.08***	0.03***
	(11.94)	(13.44)	(14.97)	(-1.35)	(3.44)	(-5.68)	(6.72)	(-6.70)	(4.31)
$Amihud\ Illiquidity_{i,t}$	0.02***	0.02**	0.02**	1.28***	0.70***	-1.03***	1.66***	0.01	0.01
	(2.72)	(2.53)	(2.33)	(5.67)	(3.35)	(-3.49)	(5.31)	(0.58)	(1.28)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,155	20,155	20,155	20,155	20,155	20,155	20,155	19,155	19,342
Adjusted R-squared	0.736	0.734	0.699	0.905	0.580	0.076	0.345	0.040	0.835

Table 7: First Stage Stock-Day IV Estimate

Table 7 presents results for the following regression equation: $Attempted\ Spoofing_{i,t} = \beta_1 Litigation_t \times Treat_i + \boldsymbol{\beta} X + \zeta_i + \gamma_t + \epsilon_{i,t}$, where $Attempted\ Spoofing_{i,t}$ is the standardized attempted spoofing order volume scaled by trading volume for stock i on day t, $Litigation_t$ is an indicator variable equal to 1 if the date t is one to three days after an SEC litigation release on market manipulation, and $Treat_i$ is an indicator variable equal to 1 if stock i is cross-listed on a U.S. exchange. X represents standardized controls for the lag of average dollar spread, lag of average price, lag of inverse price, absolute return, log of trading volume, and Amihud (2002) illiquidity. We include stock and date fixed effects with ζ_i and γ_t, respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) $Attempted\ Spoofing_{i,t}$
$US\ Listed_i \times Litigation_t$	-0.11***
	(-3.23)
$Average\ Dollar\ Spread_{i,t-1}$	0.03
	(0.63)
$Average\ Price_{i,t-1}$	-0.09
	(-0.64)
$Inverse\ Price_{i,t-1}$	-0.37***
	(-4.89)
$\ln(Dollar\ Volume)_{i,t}$	-0.11**
	(-2.00)
$Absolute\ Return_{i,t}$	-0.00
	(-0.23)
$Amihud\ Illiquidity_{i,t}$	-0.01
	(-0.75)
Stock FE	Yes
Date FE	Yes
Observations	20,155
Adjusted R-squared	0.495

Table 8: Second Stage Stock-Day IV Estimate

Table 8 presents results for the following regression equation $Market\ Quality_{i,t} = \beta_1 \widehat{Attempted\ Spoofing}_{i,t} + \boldsymbol{\beta} X + \zeta_i + \gamma_t + \epsilon_{i,t}$, where $MarketQuality_{i,t,j}$ is the log of 15-second return volatility, log of 1-minute return volatility, log of 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted price impact, log of variance ratio, or log of Hasbrouck (1993) pricing error σ. $\widehat{Attempted\ Spoofing}_{i,t}$ is the predicted standardized attempted spoofing volume scaled by trading volume for stock i on day t from the first-stage IV regression in Table 7. X represents standardized controls for the lag of average dollar spread, lag of average price, lag of inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock and date fixed effects with ζ_i and γ_t , respectively. T-statistics are reported in parentheses and standard errors are clustered by stock.

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$\widehat{Attempted\ Spoofing}_{i,t}$	0.24**	0.28**	0.27*	0.88	5.74*	-1.41	5.74	0.24	0.39*
	(1.99)	(2.29)	(1.94)	(0.53)	(1.79)	(-0.41)	(1.42)	(0.51)	(1.71)
$Average\ Dollar\ Spread_{i,t-1}$	0.09***	0.09***	0.08***	2.29***	1.07***	0.64**	0.57	0.03	0.07**
	(5.09)	(4.72)	(4.63)	(6.13)	(2.76)	(2.51)	(1.32)	(0.73)	(2.24)
$Average\ Price_{i,t-1}$	-0.08	-0.07	-0.07	-0.79	0.12	-0.44	0.56	0.00	0.18*
	(-1.64)	(-1.42)	(-1.40)	(-1.44)	(0.13)	(-0.73)	(0.53)	(0.06)	(1.72)
$Inverse\ Price_{i,t-1}$	0.15***	0.17***	0.17***	10.80***	8.38***	5.05***	3.25*	0.04	0.16
	(2.68)	(3.12)	(3.06)	(10.32)	(6.16)	(3.70)	(1.96)	(0.25)	(1.10)
$\ln(Dollar\ Volume)_{i,t}$	0.26***	0.26***	0.25***	-0.32	-2.07***	1.86***	-3.71***	0.03	-0.30***
	(9.77)	(9.19)	(8.41)	(-1.27)	(-3.77)	(2.73)	(-4.76)	(0.50)	(-6.32)
$Absolute\ Return_{i,t}$	0.06***	0.07***	0.08***	-0.09	0.42***	-0.95***	1.33***	-0.08***	0.03***
	(10.41)	(11.09)	(12.63)	(-1.25)	(3.11)	(-5.62)	(6.29)	(-6.66)	(3.79)
$Amihud\ Illiquidity_{i,t}$	0.02**	0.02**	0.02**	1.29***	0.77***	-1.05***	1.73***	0.01	0.02
	(2.44)	(2.23)	(2.08)	(5.67)	(3.23)	(-3.45)	(4.95)	(0.70)	(1.33)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,155	20,155	20,155	20,155	20,155	20,155	20,155	19,155	19,342
Kleibergen-Paap F-statistic	10.41	10.41	10.41	10.41	10.41	10.41	10.41	8.906	8.423

Table 9: Alternate Spoofing Measures

Table 9 presents results for the following regression equation: $Market\ Quality_{i,t,j} = \beta_1 \widehat{Spoofing}_{i,t,j} + \boldsymbol{\beta} X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j}$, where $MarketQuality_{i,t,j}$ is the log of 15-second return volatility, log of 1-minute return volatility, log of 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted price impact, log of variance ratio, or log of Hasbrouck (1993) pricing error σ. $\widehat{Spoofing}_{i,t,j}$ is the predicted standardized successful or unsuccessful spoofing volume scaled by trading volume for stock i on day t and 30-minute interval j using the first-stage IV regression specification in Table 4. X represents standardized $j-1$ controls for the average dollar spread, average price, inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock fixed effects with ζ_i and 30-minute interval fixed effects with $\phi_{t,j}$. T-statistics are reported in parentheses and standard errors are clustered by stock-day. Panels A and B present second-stage IV results at the intraday level for successful and unsuccessful spoofing, respectively.

Panel A: Successful Spoofing

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$\widehat{Successful\ Spoofing}_{i,t,j}$	0.33***	0.28***	0.23***	-2.86***	0.50	-1.59**	1.93***	0.20***	0.05***
	(14.13)	(12.46)	(10.00)	(-8.79)	(0.75)	(-2.18)	(3.55)	(4.98)	(2.68)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276,808	275,479	272,831	292,772	292,715	292,715	292,715	260,371	203,750
Kleibergen-Paap F-statistic	330.4	329.6	328.3	316.4	316.4	316.4	316.4	337.1	329.8

Panel B: Unsuccessful Spoofing

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$\widehat{Unsuccessful\ Spoofing}_{i,t,j}$	0.28***	0.24***	0.20***	-2.26***	0.40	-1.26**	1.52***	0.18***	0.05***
	(14.95)	(12.82)	(10.09)	(-9.24)	(0.75)	(-2.21)	(3.56)	(5.08)	(2.69)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276,808	275,479	272,831	292,772	292,715	292,715	292,715	260,371	203,750
Kleibergen-Paap F-statistic	436.2	435.6	434.3	523.7	523.8	523.8	523.8	436.4	424.3

Table 10: HFT vs. Non-HFT Spoofing

Table 10 presents results for the following regression equation: $Market\ Quality_{i,t,j} = \beta_1 \widehat{Spoofing}_{i,t,j} + \boldsymbol{\beta}X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j}$, where $MarketQuality_{i,t,j}$ is the log of 15-second return volatility, log of 1-minute return volatility, log of 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted price impact, log of variance ratio, or log of Hasbrouck (1993) pricing error σ. $\widehat{Spoofing}_{i,t,j}$ is the predicted standardized HFT or non-HFT spoofing volume scaled by trading volume for stock i on day t and 30-minute interval j using the first-stage IV regression specification in Table 4. X represents standardized $j-1$ controls for the average dollar spread, average price, inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock fixed effects with ζ_i and 30-minute interval fixed effects with $\phi_{t,j}$. T-statistics are reported in parentheses and standard errors are clustered by stock-day. Panels A and B present second-stage IV results at the intraday level for HFT and non-HFT spoofing, respectively.

Panel A: HFT Spoofing

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$HFT\ \widehat{Spoofing}_{i,t,j}$	0.20***	0.17***	0.14***	-1.76***	0.31	-0.98**	1.19***	0.12***	0.03***
	(19.08)	(16.32)	(11.97)	(-9.40)	(0.75)	(-2.23)	(3.63)	(5.06)	(2.68)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276,808	275,479	272,831	292,772	292,715	292,715	292,715	260,371	203,750
Kleibergen-Paap F-statistic	938.2	936	934.5	963.9	963.9	963.9	963.9	937.8	888.8

Panel B: Non-HFT Spoofing

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$Non - \widehat{HFT\ Spoofing}_{i,t,j}$	0.25***	0.21***	0.17***	-2.01***	0.36	-1.12**	1.36***	0.16***	0.04***
	(13.62)	(11.88)	(9.57)	(-9.04)	(0.75)	(-2.20)	(3.52)	(5.05)	(2.66)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276,808	275,479	272,831	292,772	292,715	292,715	292,715	260,371	203,750
Kleibergen-Paap F-statistic	328.2	327.9	327.7	405.8	405.8	405.8	405.8	324.9	314

Table 11: Trader-Level Filters

Table 11 presents results for the following regression equation: $Market\ Quality_{i,t,j} = \beta_1 \widehat{Spoofing}_{i,t,j} + \boldsymbol{\beta} X + \zeta_i + \phi_{t,j} + \epsilon_{i,t,j}$, where $MarketQuality_{i,t,j}$ is the log of 15-second return volatility, log of 1-minute return volatility, log of 5-minute return volatility, time-weighted quoted spread, volume-weighted effective spread, volume-weighted realized spread, volume-weighted price impact, log of variance ratio, or log of Hasbrouck (1993) pricing error σ. $\widehat{Spoofing}_{i,t,j}$ is the predicted standardized spoofing volume (with trader-level filters imposed) scaled by trading volume for stock i on day t and 30-minute interval j using the first-stage IV regression specification in Table 4. X represents standardized $j-1$ controls for the average dollar spread, average price, inverse price, log of trading volume, absolute return, and Amihud (2002) illiquidity. We include stock fixed effects with ζ_i and 30-minute interval fixed effects with $\phi_{t,j}$. T-statistics are reported in parentheses and standard errors are clustered by stock-day.

	(1) 15-second volatility	(2) 1-minute volatility	(3) 5-minute volatility	(4) Quoted spread	(5) Effective spread	(6) Realized spread	(7) Price impact	(8) Variance ratio	(9) Hasbrouck σ
$Filtered\ \widehat{Spoofing}_{i,t,j}$	0.25***	0.21***	0.17***	-2.00***	0.35	-1.11**	1.35***	0.15***	0.04***
	(15.21)	(13.10)	(10.25)	(-9.21)	(0.75)	(-2.22)	(3.57)	(5.08)	(2.69)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
30-minute FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	276,808	275,479	272,831	292,772	292,715	292,715	292,715	260,371	203,750
Kleibergen-Paap F-statistic	451.7	451.2	450.2	538.9	538.9	538.9	538.9	440.4	418.1

Internet Appendix for
Does High Frequency Market Manipulation Harm Market Quality?

Jonathan Brogaard

David Eccles School of Business,
University of Utah

Dan Li

School of Management and Economics,
Chinese University of Hong Kong, Shenzhen

Jeffrey Yang

David Eccles School of Business,
University of Utah

Appendix A1: Spoofing Court Cases

This section describes several seminal spoofing court cases. For each case, we summarize the spoofing case, describe the court's definition of spoofing, and describe additional analysis from regulators or expert witnesses.

A1.1: *United States v. Coscia* (2017)

Background

In the first spoofing case in the United States, Michael Coscia was charged and sentenced to prison for violating the anti-spoofing provision in the Commodities Exchange Act. In *United States v. Coscia* (2017), Coscia appealed, arguing that the regulations were too vague and that there was a lack of evidence that supported his guilt. The court rejected the appeal, arguing that the anti-spoofing provision was not unconstitutionally vague, and that there was sufficient evidence to convict Coscia.

Michael Coscia owned and operated the high-frequency trading firm Panther Energy Trading, LLC. In the initial trial in 2015, Coscia was accused of spoofing in the commodities futures market in 2011. To facilitate the spoofing strategy, Coscia hired Jeremiah Park to create two trading programs, Flash Trader and Quote Trader. Park's testimony revealed that the programs would bait the algorithmic traders with large spoofing orders and move prices in one direction, while at the same time profitably executing small genuine orders. The trading strategy yielded Coscia $1.4 million in a three month period. Coscia was ordered to return the $1.4 million in addition to a $1.4 million fine, and was sentenced to three years in prison.

Court's definition of spoofing (United States v. Coscia, 807 F.3d 610 (7th Cir. 2017))

> In practice, spoofing, like legitimate high-frequency trading, utilizes extremely fast trading strategies. It differs from legitimate trading, however, in that it can be employed to artificially move the market price of a stock or commodity up and down, instead of taking advantage of natural market events (as in the price arbitrage strategy discussed above). This artificial movement is accomplished in a number of ways, although it is most simply realized by placing large and small orders on opposite sides of the market. The small order is placed at a desired price, which is either above or below the current market price, depending on whether the trader wants to buy or sell. If the trader wants to buy, the price on the small batch will be lower than the market price; if the trader wants to sell, the price on the small batch will be higher. Large orders are then placed on the opposite side of the market at prices designed to shift the market toward the price at which the small order was listed. (p. 5-6)

Expert witness analysis

Professor Hank Bessembinder was an expert witness for the U.S. Government during the Coscia trial. While he did not provide a formal definition of spoofing, he provided ex-post evidence for why Coscia's trading behavior was suggestive of spoofing behavior[13]. First, the fill rates on the genuine orders were high relative to both the spoofing orders and HFT orders in general. Professor Bessembinder is quoted as saying that "there were more than 10 times as many contracts traded on the small orders as compared to the large orders." Second, Coscia had a large imbalance in large and small orders that differed from that of other HFTs. Professor Bessembinder stated that Coscia "was entering over 60 percent of his orders as large orders, whereas, the other high-frequency traders were entering only about a quarter of one percent of their orders as large orders." Lastly, Coscia had a high order cancellation rate, particularly in the large spoofing orders.

[13] The Bessembinder expert witness details are from two sources. The first point on fill rates is found in the 2016 U.S. v. Coscia case (United States v. Coscia, 177 F. Supp. 3d 1087 (N.D. Ill. 2016)), while the latter points are found in the 2021 U.S. v. Coscia case (United States v. Coscia, 4 F.4th 454 (7th Cir. 2021)).

Professor Bessembinder stated that Coscia cancelled "a little over 97 percent" of large orders within a second, while the percentages was under 35 percent for other HFTs.

A1.2: U.S. Commodity Futures Trading Commission v. Nav Sarao Futures Limited PLC and Navinder Singh Sarao (2017)

Background

Navinder Singh Sarao was accused and convicted of spoofing E-mini S&P 500 futures in 2009 to 2015. Based in his parents' home in London, Sarao used a combination of manual spoofing and algorithmic spoofing to manipulate E-mini prices. Importantly, Sarao's spoofing activity is considered one of the contributors to the May 6, 2010 Flash Crash. Sarao was arrested in 2015 and paid $12.8 million to the U.S. government.[14]

Court's definition of layering (Case No. 15-cv-3398)

> More specifically, Defendants placed thousands of orders to buy or to sell E-mini S&P futures contracts that they did not intend to execute at the time the orders were placed (Spoof Orders). Defendants' intent in placing these Spoof Orders was to create a materially false and misleading impression of supply (when placing sell-side Spoof Orders) and demand (when placing buy-side Spoof Orders) in order to induce other market participants to react to the false Spoof Order information and to buy or sell E-mini S&P futures contracts at prices, quantities, and/or times that, but for Defendants' Spoof Orders, they would not otherwise have traded. (p.6)

Regulatory analysis[15]

The regulatory complaint included several metrics that were used to argue that Sarao's layering orders were different from typical E-mini orders. First, Sarao's layering algorithm had extremely

---

[14] https://www.bbc.co.uk/news/business-37932250
[15] We use the regulatory analysis from the CFTC's complaint: https://www.cftc.gov/sites/default/files/idc/groups/public/@lrenforcementactions/documents/legalpleading/enfsaraocomplaint041715.pdf

high cancellation rates, with a cancellation rate of over 99%. However, the cancellation rate from other traders for orders of similar sizes was 49%. Second, Sarao's layering algorithm order size was 504 contracts on average, while the order size was 7 contracts on average for other traders. Third, Sarao's layering algorithm had a higher amendment rate, with 161 amendments on average per order, while orders from other traders averaged only 1 modification per order.

A1.3: Securities and Exchange Commission v. Lek Securities Corporation, et al (2019)

Background

On March 20, 2017, the SEC sued Lek Securities Corporation for market manipulation. Lek Securities is a New York-based broker-dealer firm. Avalon Securities used the Lek broker-dealer services to trade in the U.S. The SEC accused Avalon Securities of market manipulation, and argued that Lek Securities profited from Avalon's manipulative activity. The jury verdict in 2019 was in favor of the SEC, which resulted in fines for Lek and Avalon.

The jury found that the defendants were guilty of layering and cross-market spoofing. From 2012 to 2016, the defendants had over 675,000 cases of layering and 668 cases of cross-market spoofing, which yielded over $21 million from layering and over $7 million from cross-market spoofing.

Court's definition of layering (612 F. Supp. 3d 287 (S.D.N.Y. 2020))

> The first manipulative scheme, referred to as "layering," involved placing multiple orders to buy (or sell) a given stock at increasing (or decreasing) prices, to move the price of the security without intending to execute those orders. These are referred to as the loud-side orders. The loud-side orders created the appearance of an artificially inflated level of demand (or supply) for a stock. In conjunction with the loud-side orders, the trader would place a smaller number of orders on the opposite side of the market to sell (or buy) the same stock. These are referred to as the quiet-side orders. Once the stock reached the desired price, the trader canceled the loud-side orders.

Expert witness analysis (370 F. Supp. 3d 384 (S.D.N.Y. 2019))

Professors Terrence Hendershott and Neil Pearson provided expert witness testimony. Because the defendants moved to exclude the expert testimony in trial, the court described the testimonies in detail. We focus on Professor Hendershott's "layering loop" analysis. Below is an excerpt from the court opinion (p.4).

> Hendershott applied five criteria to identify groups of orders, cancellations, and executions consistent with layering. First, Hendershott considered only instances where a trader places both buy and sell orders in a single stock, because layering is a strategy that involves a trader placing orders on both sides of the market. Second, Hendershott only considered instances where the orders were entirely resolved through cancellation or execution within 60 seconds, even though it is possible for traders to engage in a layering scheme through transactions that last longer than 60 seconds. The parties refer to these groupings as "Loops."
>
> Third, Hendershott required both the number of visible orders and the number of shares in those orders on the Loud side of a Loop to be greater than both the orders and shares on the Quiet side by at least two to one (the "Order Imbalance"). Approximately 2 million Loops from the Avalon Trade Data met Hendershott's first three criteria.
>
> Fourth, Hendershott eliminated Loops where the ratio of executed shares on the Quiet side to the Loud side was less than three to one (the "Execution Imbalance"), even though the Loud-side shares were more numerous. Hendershott contends that considering only Loops with an Execution Imbalance of at least three to one eliminates trading strategies such as market making from the Loops.
>
> Fifth, Hendershott eliminated Loops if a Loud-side order was placed more than one second after the last Quiet-side execution or

> cancellation. He reasoned that this was consistent with a layering strategy, which typically involves placing Loud-side orders to achieve favorable execution prices for Quiet-side orders. Hendershott explains that, together, these five criteria create a conservative data set reflecting patterns of layering activity. Applying these criteria yielded a total of 675,504 Loops that Hendershott found to be consistent with layering (the "Layering Loops"). Of those, 663,994 occurred after March 12, 2012.

In addition to the layering definition, Hendershott's testimony provides additional analysis that suggests that Avalon's activity is layering. First, he shows that the layering orders are cancelled within seconds of the genuine order executions, which he states "is consistent with a layering strategy which tries to minimize the execution rate of Loud-side orders" (p. 4). Second, Hendershott provides evidence that suggests that Avalon's activity has no economic rationale. He shows that the realized spread associated with the genuine-side orders tends to be positive, while the realized spread associated with the layering orders tends to be negative. This is evidence that the layering orders did not have an "economic rationale." (p. 5)

A1.4: Harrington Global Opportunity Fund v. CIBC World Markets Corp

Background

Harrington Global Opportunity Fund owned shares of Concordia, a pharmaceutical company that is cross-listed on NASDAQ and TSX. From January 27, 2016 to November 15, 2016, the price of a Concordia share fell from $28.03 to $3.13. Defendants are traders from both the U.S. and Canada, whom Harrington accused of manipulative spoofing and short-selling.

Because there were defendants from both the U.S. and Canada, the court had to determine whether it had jurisdiction over the Canadian traders. The court argued that under the "effects test," it had jurisdiction over the Canadian traders, even if they did not manipulate Concordia on U.S. exchanges. This is because manipulating the price of Concordia listed on the TSX will also affect the price of shares on the NASDAQ. The court dismissed the manipulative short-selling claims in February of 2022, while the spoofing claims remained (585 F. Supp. 3d 405 (S.D.N.Y. 2022)).

Court's definition of spoofing (585 F. Supp. 3d 405 (S.D.N.Y. 2022))

> Nonetheless, when looking to indicia that distinguish spoofing from legitimate market activity, courts tend to examine (1) the passage of time between placement and canceling of orders (usually in milliseconds), (2) cancellation of orders when large baiting orders are partially filled or legitimate small orders are completely filled, (3) parking baiting orders behind smaller legitimate orders placed by other traders and (4) large disparities in the volume of baiting orders on one side of the market and legitimate orders placed by the spoofer. (p. 7)

A1.5: U.S. Commodity Futures Trading Commission v. Oystacher et al.

Background

The CFTC accused Igor Oystacher and 3Red Trading LLC of spoofing on over 51 days from December 2011 to January 2014. The defendants were accused of spoofing E-Mini S&P 500 futures, crude oil and natural gas futures, copper futures, and VIX futures.[16] The CFTC and defendants settled in 2016, resulting in a $2.5 million civil fine and trading limitations for the defendants.

Expert witness analysis (No. 15-CV-9196)

Professor Hendrik Bessembinder was hired by the CFTC as an expert witness. His analysis is detailed in the court's Memorandum Opinion and Order. To identify instances of spoofing, Bessembinder first identified "flipping patterns." In particular, he stated that "[a] flip refers to [the] cancellation of an order followed by an opposite side order entry within 0.005 seconds and at the same or better price." (p.46).

Professor Bessembinder then used the following four criteria to further narrow down the flip orders (p.46-47). Flip orders were included if they:

---

[16] https://www.cftc.gov/PressRoom/PressReleases/7264-15

1. Were “placed and cancelled in less than a second”
2. “At least doubled the quantity of contracts that was already in the limit order book at the relevant prices”
3. “Were placed at an existing price—i.e., did not establish a new best bid or offer”
4. “Were fully visible to the market—i.e., not iceberg orders”